12-1991

The development of a voice-interactive avionics system for the AV-8B Harrier II tactical jet aircraft

Dennis Patrick O'Donoghue

To the Graduate Council:

I am submitting herewith a thesis written by Dennis Patrick O'Donoghue entitled "The development of a voice-interactive avionics system for the AV-8B Harrier II tactical jet aircraft." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Aviation Systems.

Ralph D. Kimberlin, Major Professor

We have read this thesis and recommend its acceptance:

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a thesis written by Dennis Patrick O'Donoghue entitled "The Development of a Voice-Interactive Avionics System for the AV-8B Harrier II Tactical Jet Aircraft." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Aviation Systems.

Ralph D. Kimberlin, Major Professor

We have read this thesis
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor
and Dean of The Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at The University of Tennessee, Knoxville, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature Dennis Patrick O'Donoghue

Date 30 OCTOBER 1991

THE DEVELOPMENT OF A VOICE-INTERACTIVE AVIONICS SYSTEM FOR THE AV-8B HARRIER II TACTICAL JET AIRCRAFT

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Dennis Patrick O'Donoghue
December 1991

DEDICATION

To

Dr. Lewis John Guarnieri D.D.S, Ph.D

and

Dr. Nancy Stuart-Maxwell Guarnieri Ed.D.

Thank you,

for everything.

ABSTRACT

This research explored the feasibility of implementing a voice-interactive avionics system into the cockpit of current operational AV-8B Harrier tactical jet aircraft, in an attempt to reduce high pilot workload during tactical phases of flight. A review of Automatic Speech Recognition technology development and capabilities was conducted, as well as an overview of previous and on-going cockpit voice-interactive research projects. A Voice-Interactive System (VIS) for a TAV-8B testbed aircraft was developed in the Spring of 1990, and a VIS technology demonstration program was conducted throughout the following summer. A total of 12 evaluation flights were conducted by two test pilots experienced in Harrier tactical flight operations. Objective and subjective measures of VIS performance were documented through the use of cockpit recording equipment and pilot comments during flight and post-flight debriefings. Test results indicated that VIS technology demonstrated excellent potential to reduce pilot workload during critical phases of flight. VIS performance results were encouraging. The system demonstrated real-time recognition rates; recognition accuracy averaged 95.7%. A great deal of difficulty, however, was experienced with the keyword activation feature of the system, which was designed to alert the system to listen to pilot commands. (Voice activation of the VIS by the keyword was achieved only 51% of the time.) Both test pilots indicated a strong desire for a manually activated VIS switch, which would provide more reliable system activation. Overall, VIS technology demonstrated a level of performance and utility that warrants further development and implementation into the operational Harrier fleet as soon as possible.

TABLE OF CONTENTS

LISTS OF FIGURES

LIST OF ABBREVIATIONS

Advanced Fighter Technology Integration	AFTI
Angle Rate Bombing System	ARBS
Automatic Speech Recognition	ASR
Auxiliary Communications Navigation Identification Panel	ACNIP
Avionics Research And Development Activity	AVRADA
Bandpass Filtering	BPF
Central Processer Unit	CPU
Close Air Support	CAS
Coder/Decoder	CODEC
Communication	COMM
Continuous Variable Slope Deltamod	CVSD
Control Point	CP
Course-Speed-Time	C-S-T
Digital Signal Processor	DSP
Display Processor	DP
Dynamic Programming	DPG
Dynamic Time Warping	DTW
Electronic Counter-Measures	ECM
Electronic Horizontal Situation Indicator	EHSI
Forward Air Controller (Airborne)	FAC(A)
Global Positioning System	GPS
Hands-On-Throttle-And-Stick	HOTAS
Heads-Up Display	HUD
Hidden Markov Models	HMM
Inertial Navigation System	INS
Initial Point	IP
Linear-Predictive Coding	LPC
Mission Computer	MC
Multi-Function Display	MFD
Multi-Purpose Display	MPD
National Aeronautics and Space Administration	NASA
Naval Air Systems Command	NAVAIRSYSCOM

Naval Air Test Center	NATC
Navigation Forward Looking Infra-Red	NAVFLIR
Night Vision Goggles	NVG
Options Display Unit	ODU
Radar Warning Receiver	RWR
Rendevous Point	RP
Signal to Noise ratio	S/N
Stores Management System	SMS
Tactical Air Controller (Airborne)	TAC(A)
Target Of Opportunity	TOO
Time-On-Target	T-O-T
United States Marine Corps	USMC
Universal Transverse Mercator	UTM
Up-Front Controller	UPC
Visual Meteorological Conditions	VMC
Voice Interactive Communications	VIC
Voice Interactive System	VIS
Voice Interface Box	VIB
Voice Calibration	VCAL
Voice Recognition System	VRS

I. INTRODUCTION

The integration of electronics and computer technology into tactical aircraft has led to avionics, flight control and weapons systems with tremendous capabilities. Indeed, modern day jet fighters have demonstrated performance that would have been deemed impossible just 20 years ago. However, the complexity of enemy air defense systems has at the same time forced the development of increasingly complex tactics to successfully evade detection or engage and defeat the threat. Long range search radars have forced attacking aircraft to fly as low as 100 feet above ground level at high speeds for extended periods of time to avoid detection. Most recently, air forces worldwide are being called upon to conduct night low-altitude ground attack missions in support of armies which are increasingly opting to attack at night under the cover of darkness. Radar, Infra-Red and optical detections systems,as well as weapons delivery systems are constantly being developed and updated to deal with the increasingly sophisticated and demanding combat environment. Paralleling this effort, human factors engineers are working to develop cockpit controls and displays in an effort to ease the workload of the single-seat pilot who must operate all these systems in the demanding combat environment.

Early in the decade of the 1980s the U.S. Marine Corps directed that its AV-8B Harrier day attack aircraft be reconfigured for the night attack mission. Major components of the reconfiguration consisted of infra-red sensors, special cockpit lighting, the addition of a second multi-purpose display, and the relocation of critical systems controls to the control stick and throttle. The first AV-8B "Night Attack Harrier" squadron became operational in late 1989. Night attack pilots were quick to point out that while the new systems enabled them to successfully complete the mission, the demands placed upon them to operate those systems while executing low-altitude tactical maneuvers was resulting in extremely high pilot workload.

Cockpit designers are maximizing the use of the pilot's visual auditory and tactile senses in the operation of cockpit controls and displays. Warning tones and synthesized voice warnings sound to alert the pilot of vital systems malfunctions. Control switches are designed with distinctive shapes to enable the pilot to identify controls by tactile cues alone. Heads-Up displays, multi-purpose displays and glass instrument panels have been designed to optimize critical display space. However, while the AV-8B cockpit is unarguably at the leading edge of ergonomic design, test pilots continue to uncover mission-oriented flight tasks where pilot workload is prohibitively high.

The one mode of control not exploited by cockpit designers is speech. Speech is a very natural human mode of interaction. Enabling the pilot to control a weapons system or activate a switch or change a display by voice command would significantly reduce pilot workload. Of even more benefit would be a system that checks and reports on systems status after a verbal request from the pilot. Pilots would truly be able to remain 'head out of the cockpit, hands on the stick and throttle' and devote more attention to target location and terrain avoidance. Mission effectiveness and flight safety would be significantly increased.

Historically, the application of speech technology to the cockpit has been limited to speech synthesis of voice warnings. Although speech recognition systems have been developed and successfully demonstrated as early as 1958, the cockpit environment has presented special challenges that have proved difficult to overcome. Recognition of voice commands in the presence of high levels of background noise, and the development of natural and flexible vocabularies have been the central focus of cockpit voice research. To date, the integration of voice recognition systems into military aircraft has been limited to research and development programs. However, state-of-the-art speech recognizers have matured to the point that the integration of a voice interactive system into AV-8B Harrier aircraft for operational use is now feasible.

In the summer of 1990 the Naval Air Test Center conducted a test and evaluation program of an ITT voice recognition system in a two-seat AV-8B testbed aircraft. The project yielded some encouraging results as well as a few substantial deficiencies. This paper reviews the results of that project and explores possible solutions to the problems encountered. The first part of this paper will reviews the evolution of the Harrier cockpit workstation and the associated increase in pilot workload; as well as the development of voice recognition systems and the application of that technology to military aircraft in earlier research programs. The second part details the development and integration of a voice interactive cockpit for the AV-8B aircraft and discusses the results of the test program. Finally, recommendations are proposed for further development and implementation of voice interactive technology for the Harrier cockpit.

II. BACKGROUND

The AV-8B Harrier Night Attack, the latest in the Harrier series of aircraft, is a single-cockpit, single-engine tactical jet aircraft, built by the McDonnell-Douglas Corporation. The aircraft is designed for the day and night visual ground attack mission, and features a Navigation Forward Looking Infra-Red (NAVFLIR) system and a Night Vision Goggle (NVG) compatible cockpit for the night attack missions; a passive Angle Rate Bombing System (ARBS), a laser spot tracker for target designation, and an inertially-aided weapons delivery system. The aircraft is scheduled to receive the APG-65 multi-mode radar system in early 1992. A three view drawing of the aircraft is shown in figure 1.

The evolution of the AV-8 Harrier series tactical jet aircraft has seen dramatic changes in cockpit hardware and design. The AV-8A, the first aircraft in the series, was introduced into military service in 1968 with the U.S. Marine Corps. The cockpit design, shown in figure 2, reflects the typical avionics and weapons systems controls and displays of 1960's generation tactical aircraft. It should be pointed out that,even in its day, the AV-8A was a relatively unsophisticated day attack aircraft. It possessed no radar system and only one air-ground ordnance delivery mode. Consequently, all essential aircraft systems controls

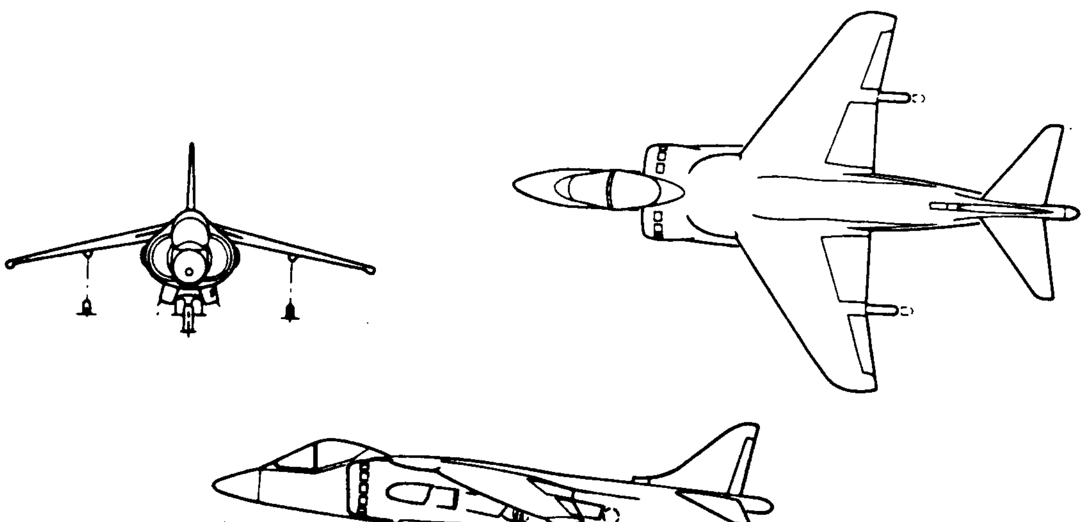


Figure 1
AV-8B HARRIER NIGHT ATTACK AIRCRAFT
Source: McDonnell-Douglas Corporation.

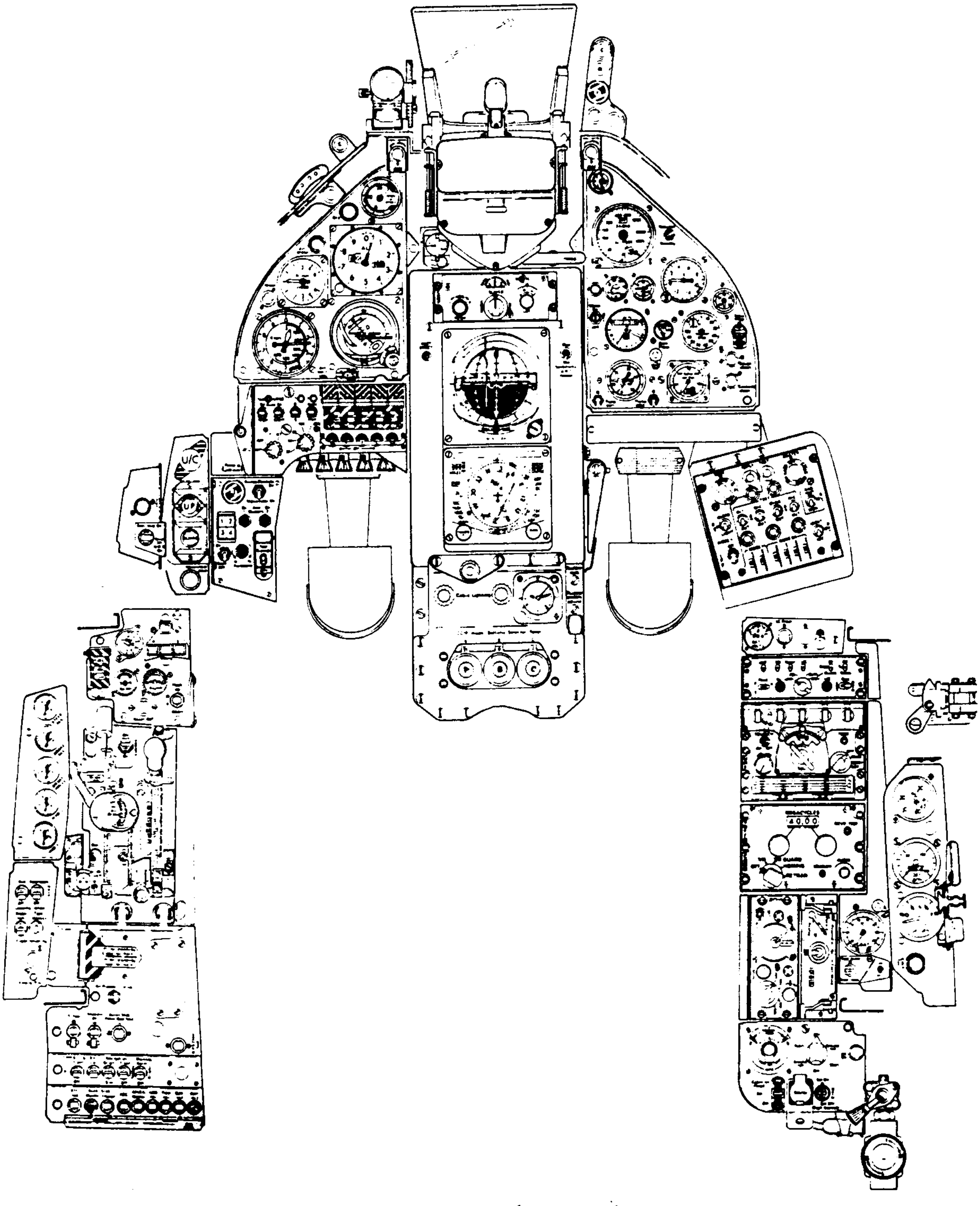


Figure 2
AV-8A COCKPIT WORKSTATION
Source: British Aerospace Corporation.

and displays were arranged in a somewhat uncluttered manner and vital aircraft information was easily obtainable through conventional analog gauges located on the instrument panel in front of the pilot. Additionally, the aircraft featured a Heads-Up Display (HUD) which displayed airspeed, altitude, angle-of-attack, heading and attitude information. This device allowed the pilot to keep his head up out of the cockpit and still have critical aircraft information readily available. Pilot workload was significantly reduced as pilots found they could devote more time to scanning outside the aircraft for target acquisition and terrain avoidance tasks. The development of the HUD was the first major step in the design effort to reduce pilot workload and marked the departure from the conventional "steam gauge" cockpits of the previous 50 years.

The next aircraft in the Harrier family was the AV-8B Harrier II, which was introduced into Marine Corps service in late 1983. Developed in the late 1970s as a major upgrade to the AV-8A, the aircraft featured state-of-the-art avionics and weapons systems which significantly increased combat capability, although the aircraft was still limited to the day attack mission due to a lack of night navigation and targeting sensors. The cockpit, depicted in figure 3, reflects the revolution in cockpit design that occurred in the decade of the 70s and indicates the strong influence of human factors engineering in pilot workstation design. With the exception of the small backup flight instruments on the center console, the analog gauges are gone. In their place are digital engine and fuel indicator panels; an improved field of view HUD; a Multi-Purpose Display (MPD) for display of navigation, weapons and targeting information; and an Up Front Controller (UPC) for data entry, radio, transponder and navigation systems control. The MPD, which is a conventional monochromatic cathode ray tube display surrounded by 20 pushbuttons, greatly improved the amount of data that could be displayed as compared to previous aircraft. The pilot had simply to access the menu page, then select the appropriate display page from the over 60 pages available. The amount of data and pages available to the pilot were limited only by the software capacity of the aircraft's mission computer.

In keeping with current design philosophy, the cockpit was designed around the "Hands-On-Throttle-And-Stick" (HOTAS) concept. The HOTAS concept was intended to reduce the pilot's workload by placing critical weapons systems controls at his finger tips, i.e., on the throttle and control stick so that it would not be necessary for the pilot to look inside the cockpit to locate a control or take his hands off the stick or throttle to operate that control. This design feature proved to be particularly valuable during air combat where keeping sight of the enemy aircraft is paramount, and more importantly, during low-altitude target attack when looking inside the cockpit diverts the pilots attention from the critical task

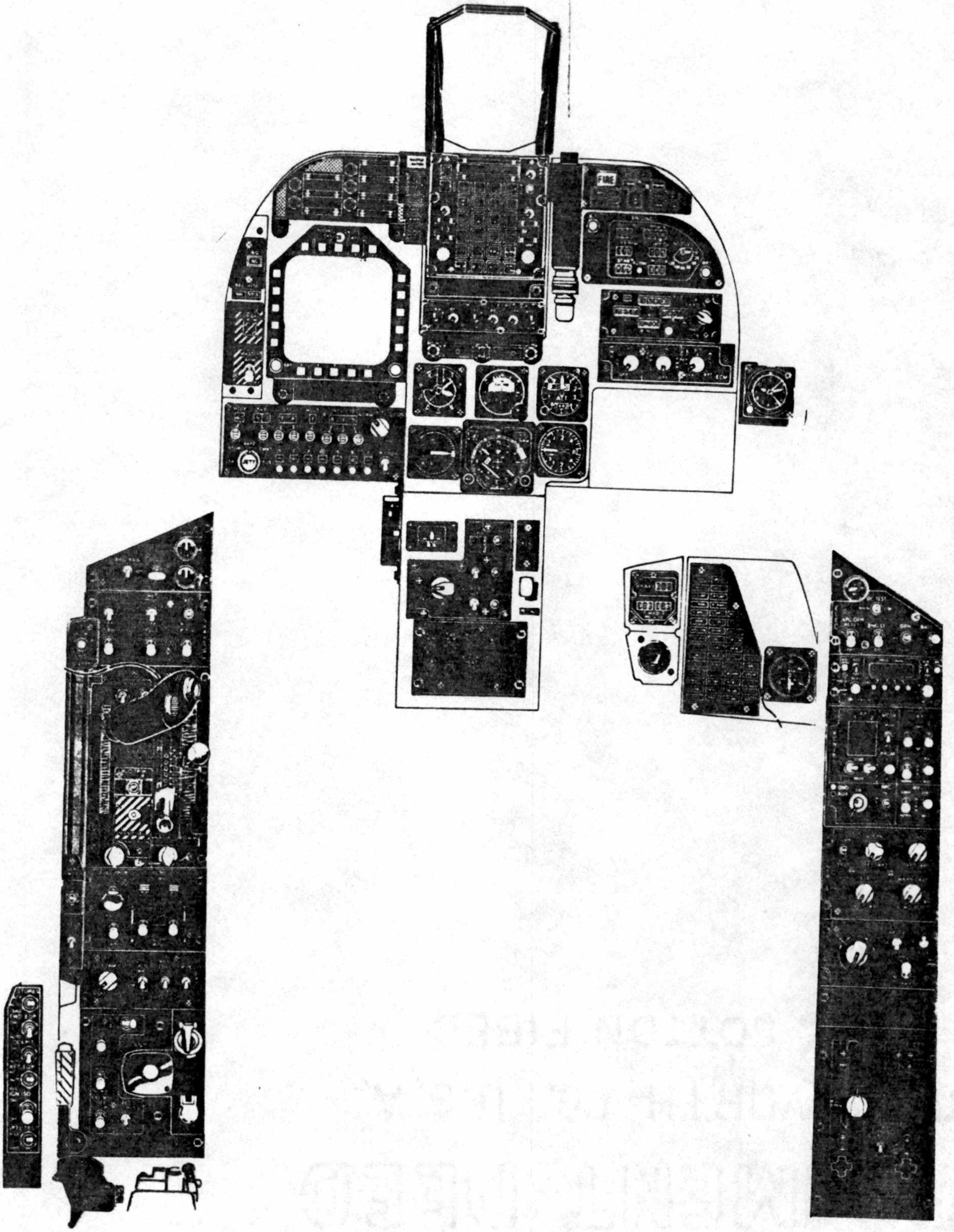


Figure 3
AV-8B COCKPIT WORKSTATION
Source: McDonnell-Douglas Corporation.

of terrain avoidance. New AV-8B pilots found they could easily control and select weapons quickly without taking their eyes off the target. Without exception, pilots noted a significant increase in combat capability and tactical effectiveness upon transitioning from the AV-8A to the AV-8B.

The AV-8B can be considered a truly 'heads-up' aircraft, in that all the performance parameters necessary to fly the aircraft are presented on the HUD. Although this has greatly reduced pilot workload it has introduced some new problems. One major problem is that the vast amount of data being displayed on the HUD has led pilots to subconsciously filter out information they don't need at the time. In other words, the pilot doesn't see what he doesn't need. The data that the pilot sees or doesn't see changes depending on the task. The adverse side effect of this subconscious 'filtering' process is that there have been cases of pilots not seeing what they did not expect to see. For example, the weapons system incorporates a feature that will flash the word "safe" on the HUD if the master armament switch is not turned on during a weapons delivery. Because the weapons system will be inhibited from releasing bombs when the master arm switch is not turned on, this HUD feature was designed as a reminder to the pilot to turn the switch on when the air to ground master mode is selected. There have been numerous cases of pilots rolling in on a target and pushing the bomb release button only to realize too late that the master arm switch was not turned on and the bombs were still on the aircraft. Upon returning to base (with a full load of unexpended ordnance) the majority of these pilots emphatically stated that they never saw a flashing "safe" cue in the HUD. Many believed that somehow the system had malfunctioned and the cue just wasn't there. However, upon review of the HUD videotape, in all cases, the flashing "safe" cue was in fact present.

How could so many pilots not see something as obvious as a word flashing right in front of their eyes? It's not because these pilots were not paying attention. A more plausible explanation is that, with so much data being presented on the HUD, the pilot's visual channel is being 'over-saturated'. The filtering problem, i.e., not seeing what one does not expect to see, is a consequence of the pilot's subconscious effort to filter data and declutter the visual channel.[1]

Sensory overload was not limited to the visual sense. Pilots frequently complained of too many warning tones in the cockpit. Cockpit designers had attempted to provide distinuishable tones at various frequencies to warn the pilot of system malfunctions,

---

[1] For an interesting discussion on Heads-Up Displays and visual channel capacity see Roscoe (1989).

ground proximity warnings and aircraft departure-prone flight conditions. Each warning was assigned a unique frequency. However, the major complaint among pilots was that the tones were too numerous and in many cases indistinuishable. As a result, the pilots found themselves still looking inside the cockpit for a warning light to verify the malfunctioning system.

To alleviate this situation, in late 1986 a voice warning system was integrated into the warning and caution advisory system. The intent was to give the pilot a clear indication of a system malfunction and the severity of that malfunction without the pilot having to look inside the cockpit at the warning/caution advisory panel.. Additionally, the system would give an altitude warning when the aircraft descended below a preset altitude. The system was enthusiastically received by AV-8B pilots and significantly alleviated visual channel saturation. The integration of a voice warning system marked the first use of synthesized speech in the AV-8B cockpit and was an early indication to Harrier pilots of the potential benefits of a speech interactive system to reduce pilot workload.

Throughout the evolution of the Harrier series the number of manual control tasks continued to increase. More advanced weapons systems meant more control switches and buttons and higher manual workload to operate those systems. Not all of the controls could be integrated onto the stick and throttle, consequently the pilot was forced to look inside the cockpit to locate and operate many of the required controls. This was particularly true in the case of the MPD. While the MPD greatly increased the amount of data available to the pilot, the size of the display limited the amount of information that could be displayed simultaneously. For instance, navigation data was presented on the Electronic Horizontal Situation Indicator (EHSI) page; weapons data was presented on the Stores Management System (SMS) page; and threat radar information was contained on the Electronic Counter-Measures (ECM) page. The single MPD meant that only one data page could be displayed at a time. During high threat ground attack missions pilots were required to frequently switch back and forth between display pages to keep abreast of navigation, weapons and threat information. Changing displays required the pilot to look down inside the cockpit and manually select the proper pushbuttons on the MPD. (Since the 20 pushbuttons on the MPD are identically shaped, it proved to be almost impossible to locate the proper pushbutton by tactile cues alone.) It soon became obvious that additional display space was needed to reduce manual workload and increase the amount of data readily available to the pilot.

The development of the AV-8B Harrier 'Night Attack' cockpit, shown in figure 4, included a second MPD and doubled the display space and amount of data instantaneously available to the pilot. However, the addition of the NAVFLIR system increased the number of data pages the pilot was required to monitor and the number of controls he was required to manually operate. The net result was a further increase in cockpit workload, especially during the night low-altitude attack mission where the pilot was required to operate the NAVFLIR, weapons system, navigation system, ECM and communications system, while wearing a cumbersome NVG device attached to his helmet and placed directly in front of his eyes. As AV-8B test pilots had predicated early in the evaluation of the Harrier Night Attack aircraft, sensory overload of the pilot's visual, tactile and auditory senses was a major obstacle to successful execution of the night attack mission.[2] New methods and modes of interface were needed to enable the pilot to operate aircraft systems within the constraints of a tolerable cockpit workload.

[2] Eason (1987)

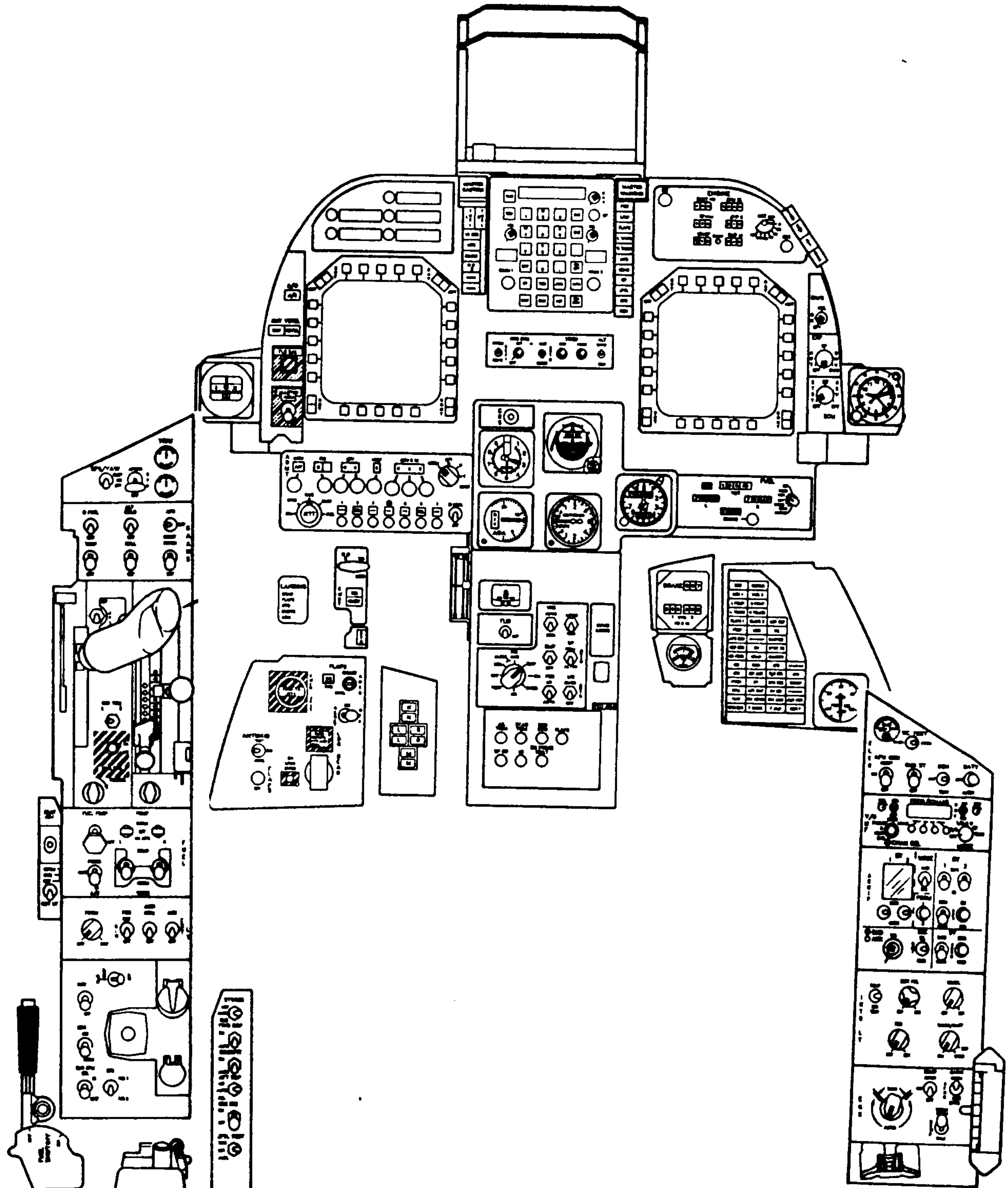


Figure 4
AV-8B NIGHT ATTACK WORKSTATION
Source: McDonnell Douglas Corporation.

III. LITERATURE REVIEW

INTRODUCTION

The critical technology in the development of a voice-interactive cockpit is Automatic Speech Recognition (ASR). To be useful, a cockpit speech recognizer must be capable of handling a reasonably flexible and natural vocabulary with a high degree of accuracy, perform in a noisy environment,.and compensate for human factors such as the effects of stress, fatigue and aircraft acceleration (g-force) on pilot speech production. Although ASR technology was successfully demonstrated as early as 1958, the vocabulary was very limited (less than 50 words), a distinct pause was required between words, and the environment was tightly controlled (low ambient noise).[3] Consequently, the constraints imposed by early "isolated-word" recognizer technology made applications to tactical aircraft impractical. Since that time however, the technology has steadily evolved to the point where medium vocabulary, continuous speech ASR systems for tactical aircraft are feasible and offer advantages over conventional modes of pilot-aircraft interface.

Communications between humans and machines has been complicated by a number of factors.[4] Major factors include the fact that various speakers pronounce words differently. Also, continuous speech sound is not easily segmented for analysis and comparison to stored reference patterns. Large vocabularies require machines to have high computational rates. And spectral similarities between different sounds require machine decisions based upon probabilities of occurrence in different speech contexts. The extent to which each of these factors will influence recognizer design is dependent on the intended use of the system. For example, a simple straight-forward task such as voice-entry of zip code data for mail sorting requires a vocabulary of only ten words (zero through nine), however, the system would necessarily be required to be speaker-independent so that many users would be able to operate the system. Additionally, such a system would have to be capable of handling rapid speech, otherwise the system would offer no advantage over entering the data manually. On the other hand, a language translator would require a very large vocabulary, say roughly 50,000 words, but it would not be unreasonable to require the speaker to say each word discretely with a distinct pause between words. A system such as this would be required to discriminate between a great many similar sounding words or

[3] Rabiner and Schafer (1978)

[4] Levinson (1990)

phrases.[5] Another system designed for medical robotics applications would be used by only a few operators (the medical team) and have a moderate, unambiguous vocabulary of 100-150 words. An acceptable and economical user interface for this type of application would be a speaker-dependent, isolated word recognizer.

The literature identifies five major dimensions that affect user interface and system design: connectedness of speech, speaker dependence, vocabulary, grammar, and environment.[6] The degree of speech connectedness that designers can attain is heavily dependent on the vocabulary and grammar required for the task, the degree of speaker-independence required, the computational capacity of the recognizer, the operational environment (background noise, operator stress, etc.), and the recognition accuracy required. Generally, it has been found that recognition accuracy is inversely proportional to both the vocabulary size and the degree of speaker-independence.[7] Although truly large vocabulary, speaker-independent, unconstrained-grammar, continuous speech recognizers, capable of operation in all types of environments, are still years away, researchers are making impressive strides towards this goal through the use of such methods as dynamic programming, active noise cancellation, finite state grammar networks and careful selection of working vocabularies.

As was quickly evident from a review of the literature, recognizer design is an optimization process with respect to two major groups of parameters: 1) operating environment, human operator limitations, and the computational and signal processing capacity of the available hardware, and 2) the requirements of vocabulary size, speech input rate, speaker-independence and recognition accuracy.

SPEECH RECOGNITION TECHNOLOGY

Currently, ASR technology is finding widespread application in the fields of manufacturing, product inspection, inventory, material handling and many other tasks where voice entry of data leaves the operators hands free to conduct other tasks.[8] However, despite the considerable success of recognizers in the commercial sector, speech recognition continues to present formidable challenges that make it a major research subject.

---

[5] Consider the similarity between the phrase "great ape" and "grey tape". Homonyms are also a major problem. For example: "to", "too", and "two".

[6] Bennett, et al (1989)

[7] Peacocke, et al (1990)

[8] See Rash (1989) and Levinson (1990) for an overview of current commercial and industrial speech recognition applications. Peacocke et al. (1990) contains an excellent performance comparison of current commercially available speech recognition systems.

The early approach to the speech recognition task was based on the apparently reasonable assumption that speech was a highly redundant signal consisting of a sequence of invariant information-bearing elements called phonemes.[9] Phonemes are the smallest units of speech that distinguish one utterance from another. (There are 40 distinct phonemes in the english language). Based on this assumption, the classical recognizer, depicted in figure 5, took the form of: a pre-processor to reduce the quantity of input data while retaining relevant information, a feature extractor to identify formant frequencies[10], a segmentor to divide the signal into phonemic segments, and a classifier to recognize individual phonemes from their features.

Figure 5
TYPICAL STRUCTURE OF AN EARLY SPEECH RECOGNIZER

Source: Moore, Roger K. (1985). "Systems for Isolated and Connected Word Recognition." In New Systems and Architectures for Automatic Speech Recognition and Synthesis. Ed. Renato DeMori and Ching Y. Suen. Berlin: Springer-Verlag. p. 75.

After the phonemes were classified, the recognition task became simply a matter of looking up the sequence of recognized phonemes in a phonemic dictionary. This approach was popular in the 1950's and 60's, however it failed to varying degrees because of the inadequacy of the initial assumption.[11] Speech signals are highly variable. One person's voice can be very different from another's, due to such factors as age, sex, or accent. Even the same speaker's voice will produce variations in the speech signal for a given word depending on whether the word is spoken softly or loudly, or when the word is spoken in

[9] Moore, Roger K. (1985)

[10] A formant frequency is a specific frequency which excites the vocal tract for a particular speech sound.

[11] Moore, R.K.(1985)

different contexts. Continuity of speech is another source of variability. Since words flow smoothly one into another, the beginnings and endings of words can change significantly. "Bread and butter" may become "breb'm butter" if spoken quickly. The literature refers to this phenomenon as "co-articulation".[12]

The co-articulation problem could be easily dealt with by requiring speakers to pronounce each word in isolation, i.e. with a distinct pause between words. However, the variability of speech signals presented researchers with a greater challenge. The problem could be partially solved by training the system to the voice characteristics of the speaker. Hence the term "speaker-dependent." This technique was most commonly accomplished by recording the speech patterns of a given speaker for a given set of words. These word patterns or "templates" were then stored for comparison during the recognition process. A block diagram of a typical pattern-matching speech recognizer is illustrated in figure 6.

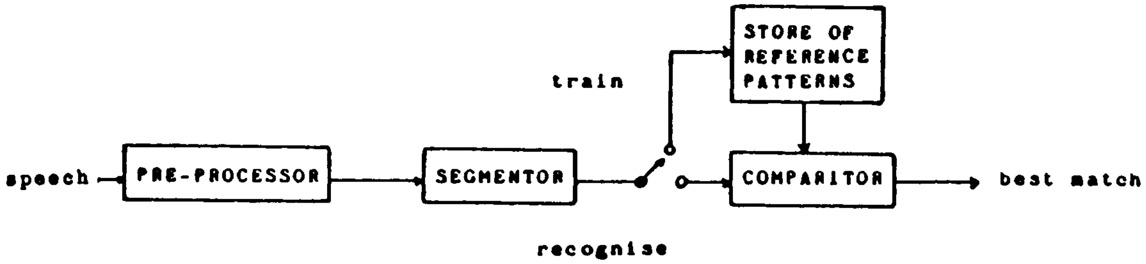


Figure 6
PATTERN-MATCHING SPEECH RECOGNIZER

Source: Moore, Roger K. (1985). "Systems for Isolated and Connected Word Recognition." In New Systems and Architectures for Automatic Speech Recognition and Synthesis. Ed. Renato DeMori and Ching Y. Suen. Berlin: Springer-Verlag. p. 80.

This type of recognizer worked fairly well as long as the speaker uttered a given word in exactly the same manner each time, however, even minor variations in speech could cause the system to mis-recognize the word.[13] What was required was a recognition algorithm that was capable of dealing with pattern similarities rather than relying on the preservation of absolute identity. Additionally, improvements in signal processing were needed to better identify and discriminate between the key characteristics of different speech signals.

---

[12] Bush and Hata (1983)
[13] Teja et al. (1983)

SIGNAL PROCESSING

The goal of signal processing is to separate speech from non-speech, perform endpoint (word boundary) detection, convert the raw waveform into a frequency domain representation, and extract and enhance only those components of the spectral representation that will be useful for the recognition task.[14]

A wide range of signal processing techniques have been applied to speech processing, however the literature identifies only a few that have established themselves as standard techniques for ASR.[15] Each of these techniques is briefly discussed below.

Short-Time Spectrum

The most common way to analyze a speech signal is to measure its short-time spectrum. This technique is based on the generally accepted assumption that a speech signal can be considered stationary over a short time interval and therefore its spectrum can be estimated by a Fourier transform analysis.[16] One of the simplest methods of short-time fourier analysis is through the use of bandpass filtering. Filter bank analyzers are easy to construct with analog circuits and the distribution of frequency bands can be readily modelled on the critical bands of the human ear. The drawback to this approach is that it is difficult to accurately estimate the spectrum shape around the spectral peaks unless a very large number of filters are used.[17] A more useful representation of the speech signal would be its wide-band spectrum.

Cepstral Analysis

The major disadvantage of wide-band spectral analysis is that it requires a short data time window. However, an alternative method, which is able to use a wider time window, is cepstral processing. This process is based on the assumption that speech is a convolution of an excitation function with a vocal tract impulse response. As illustrated in figure 7, these two components are separated by filtering the cepstrum to obtain a smooth spectrum. Usually, the first few terms of a cosine transform of the short-time log power spectrum are used.[18]

---

[14] By retaining only those parameters that are useful for recognition purposes, the amount of data that the recognition algorithm must contend with is greatly reduced.

[15] Schafer and Rabiner (1975), Flanagan (1972) and Holmes (1982).

[16] Moore, R.K. (1985).

[17] Holmes (1980)

[18] Hunt et al. (1980).

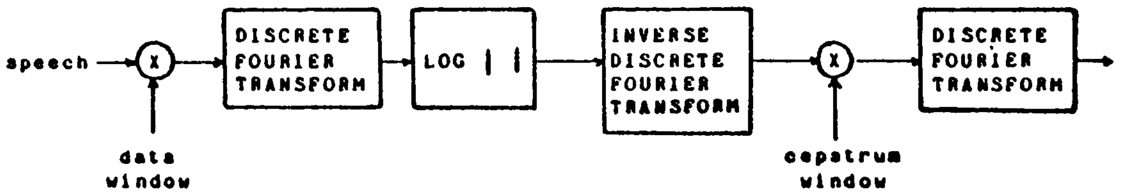


Figure 7
CEPSTRAL PROCESSING

Linear-Predictive Coding

An entirely different method of speech signal processing is Linear-Predictive Coding (LPC). Whereas short-time and cepstral processing techniques model the auditory tract, i.e., bandpass filters that model the frequency bands of the human ear, LPC models the vocal tract. This technique utilizes the autocorrelation characteristics of the speech signal and estimates the value of the current sample using a linear combination of the past n samples. From a computational point of view, this technique is very attractive. Additionally, most speech sounds correlate very well with the mathematical properties of the LPC model, thereby allowing accurate estimations of spectral peaks to be made. The exceptions to this are nasals[19] and some consonants. In these cases, LPC tends to overestimate the bandwidths of the spectral peaks, introducing some distortion in the speech signal representation.[20]

Data Reduction Techniques

The output of the signal processor is a window of data, commonly referred to as a 'speech frame'. Typically, a speech frame consists of a sequence of vectors of a given size occurring a certain number of times in a specified time interval, usually 10-30 milliseconds. Generally, this data rate is too high for recognition algorithms to handle, so various data reduction techniques have been developed. The most popular techniques are vector quantization, trace segmentation, and variable rate coding.

Vector quantization is a feature extraction technique in which the speech features F(t) in each frame are compared against a codebook of n characteristic multidimensional feature

---

19 (m,n) are examples of nasal sounds

20 Moore, R.K. (1985).

vectors {F(1), F(2), F(3)....F(n)} until the closest match Fi(t) is found. The codebook entry i(t) becomes the defacto feature for subsequent processing.[21] Although this technique does introduce some distortion into the signal representation, it is convenient and easy to implement, and significantly reduces the amount of data for subsequent signal processing.

Trace segmentation is a 'data-adaptive frame rate' technique for reducing the number of vectors in a given sequence. Each vector is considered a point in n-dimensional space, n being the size of the vector, and the trace is the sequence of points drawn out by an utterance. The total length of the trace is calculated and then divided into a fixed number of uniformly spaced intervals. The trace is then sampled. By suitable choice of the sampling rate, fewer vectors are required to represent the speech signal.[22] The advantage of this technique is that the vectors are better distributed. For example, more vectors will be present where the spectrum is changing rapidly then where it is changing slowly.

A technique very similar to trace segmentation is variable rate coding. Both techniques employ a resampling of the speech signal based on the changes in the spectrum, however, the trace segmentation technique relies on being able to determine the endpoints of the trace. In some cases this cannot be determined (such as for continuous speech), and the variable rate coding technique must be utilized.

In this technique, a threshold is set such that a vector in a sequence is retained only if the difference between its value and the value of the previously retained vector exceeds the threshold value. By adjusting the value of the threshold, the resampling rate can be adjusted to suit the requirements of subsequent processing.[23] For example, the higher the threshold, the fewer vectors or samples there will be in the more stationary regions of the spectrum.

It is important to note that all signal processing techniques unavoidably introduce some distortion into the speech signal representation. The system designer must determine how much distortion is acceptable in the design of an ASR system. Often this determination is based on the size of the selected vocabulary, and more importantly, on the phonetic similarity of the words that make up that vocabulary. Another factor that will have a major impact on the output of the signal processor is the signal-to-noise ratio (S/N) of the

---

[21] Levinson and Roe (1990).

[22] Moore, R.K. (1985)

[23] Bridle and Brown (1982) outlines in greater detail this and more complex schemes for this type of processing.

incoming signal. It is worth pointing out that, while bandpass filtering will eliminate noise above and below the bandwidth of the speech channel, it by nature cannot eliminate the noise contained within the bandwidth of the speech channel. Therefore, during the recognition process, the speech recognizer may have to contend with a significant amount of distortion and noise. Once again, wise selection of the working vocabulary can greatly minimize the impact these factors will have on recognizer accuracy.

In summary then, speech signal processing technology has been highly successful in reducing the amount of data the recognizer must process (at the cost of some signal distortion), but has been somewhat less successful in filtering noise from the signal. Other techniques have been developed to deal with the noise problem, and will be discussed later in this chapter.

THE RECOGNITION PROCESS

Speech recognition is fundamentally a pattern classification task. The literature identifies four major approaches to the pattern classification problem: pattern matching (template matching with dynamic programming), Hidden Markov Modeling (HMM), neural networks, and knowledge-based expert systems. Although neural networks and knowledge-based expert systems have been applied to speech recognition with some success, this work is still in the early stages of research, and the techniques being developed do not as yet lend themselves well to the unique requirements of a cockpit ASR system.. A detailed discussion of these relatively sophisticated techniques would be outside the scope of this thesis.[24]

Pattern Matching

The pattern matching approach to ASR was the first to receive serious attention from researchers. However, as was previously mentioned, early speech recognizers relied on the preservation of the absolute pattern identity of an utterance and even small variations in the way a speaker uttered a word had a significant impact on the performance of the recognizer. For example, if a speaker stretched out the word *hello* to *heellooo*, the recognizer would not recognize the utterance as the word *hello*. To deal with this non-linear time distortion of speech, a Dynamic Programming (DPG) technique called Dynamic

---

[24] See White (1990), Kurzweil (1989),Mariani (1989),O'Shaughnessy (1987) and Haton (1985) for an excellent discussion of recent developments in neural, expert, and knowledge-based ASR systems

Time Warping (DTW) was developed.[25] Simply stated, DTW achieves the best possible time alignment between the input signal and the reference template. The process is illustrated in figure 8. In this example the input pattern (*heellooo*), uttered from time N(1) to N(3) is mapped onto the reference pattern (*hello*) from time M(1) to M(2).

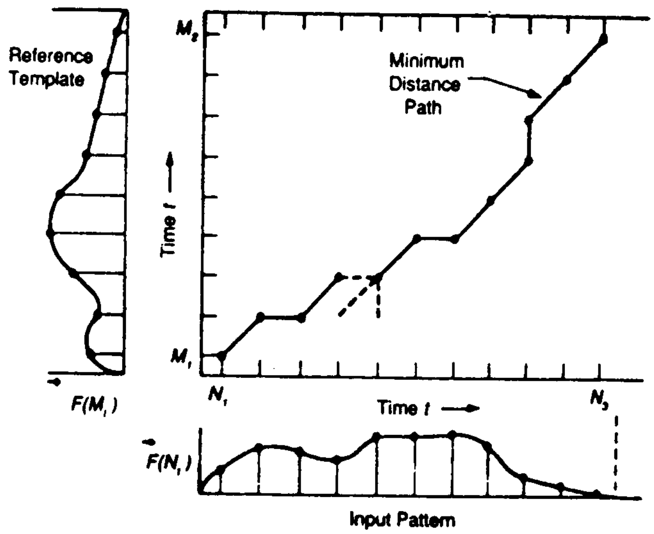


Figure 8
DYNAMIC TIME WARPING WITH SPEECH TEMPLATES
Source: Levinson, S.E. and David B.Roe (1990). "A Perspective on Speech Recognition." IEEE Communications Magazine, Jan 1990, p. 30.

Ideally, if the input and reference patterns were identical, the time alignment path would be a straight diagonal line. This rarely happens, and instead of a straight line path through the pattern, a 'minimum distance path' is mapped that minimizes the summed distance along the path. The pattern match is then based on the lowest distance or 'score' of the utterance against a set of reference templates. Employing this method, a recognizer is able to recognize a word with a speech pattern very similar to the reference template, while not absolutely identical to it. The literature indicates that DTW has been developed into a powerful DPG technique and is finding widespread use in a number of ASR applications.

---

[25] Lee (1989) provides a comprehensive and indepth discussion of a number of dynamic programming techniques.

Another DPG technique that has been highly successful in the connected word recognition task is 'level building'. In this case, an unknown utterance corresponds to a string of words and the DPG iteration illustrated in the DTW example above, must be extended so that multiple patterns can be combined to give the best match to the string. An example of the level building technique for a string of L words, is illustrated in figure 9.

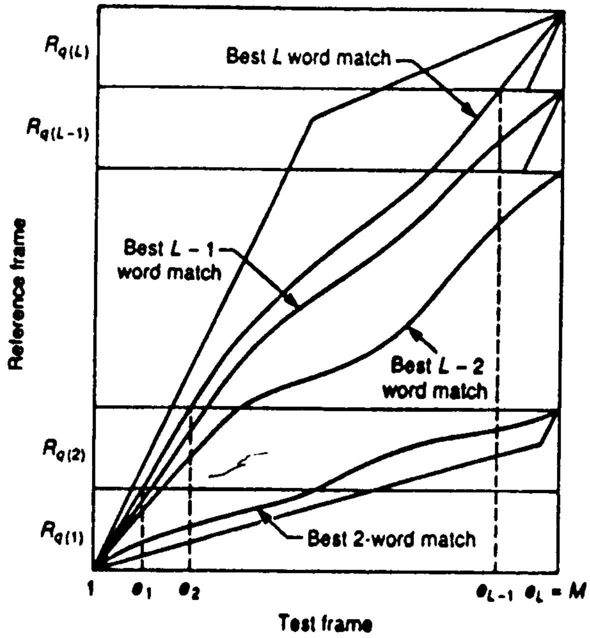


Figure 9
LEVEL BUILDING FOR CONNECTED WORD RECOGNITION
Source: Lee, C.H. (1989). "Applications of Dynamic Programming to Speech and Language Processing." AT&T Technical Journal, 68, p.119.

The goal in this technique is to obtain the optimal warping paths for strings of up to L words. The frame index of the utterance is plotted along the horizontal axis, and the vertical axis contains all possible concatenations of reference patterns that could possibly form a string of up to L words. For example, $q(l)$ is the reference pattern of the word in the l th position of a string $Rq(l)$, where $Rq(l)$ is defined as the set of all reference patterns strings of length l.. (The concatenation of reference patterns, $\{Rq(l), 1 \leq l \leq L\}$, forms a string of L words.) The horizontal lines represent the boundaries between two reference patterns in consecutive levels. The intersections of the warping paths and the horizontal lines represent the word segmentation boundaries in the utterance for the corresponding warping path. In this example, the set $\{e(l), 1 \leq l \leq L\}$ corresponds to word boundaries for the best L word match.

Although level building has become a very useful tool for connected speech recognition, its major drawback is the large computational capacity required of the recognizer or host computer. This is because the level building technique is basically a synchronous DPG search procedure, which requires that all the distance pairs for the DPG iterations be accessible for computation at any level at any instant in time.[26] Obviously, this could pose a significant computer memory management problem for large recognition tasks or real-time recognition of connected speech.

An alternative DPG technique that is more computationally efficient involves mapping the level (or grammar) information onto a finite state decision (or grammar) network.[27] In this technique, the connected word recognition problem is formulated as an optimal path searching problem through a task-oriented network, such as that shown in figure 10.

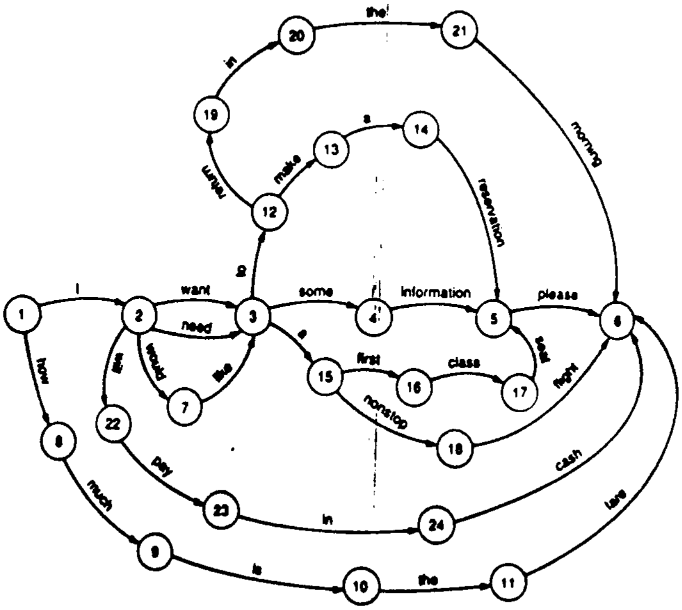


Figure 10.
TASK-ORIENTED FINITE STATE NETWORK: (A Sub-network of the AT&T Bell Laboratories Airline Reservation System Task)
Source: Levinson, S.E. and K.L. Shipley (1980). "A Conversational Mode Airline Information and Reservation System Using Speech Input and Output." Bell System Technical Journal, 59, p.121.

---

26 Lee (1989).

27 Baker (1975)

As compared to level building, which searches on a level-by-level basis, this technique maintains track of the optimal path to any node of the finite state network at every instance in time, so that the best path is obtained on a frame-by-frame basis.[28] Since only the best path information at the previous frame plus the minimum distance match (or score) at the current frame is required for the DPG iterations, this technique has the advantage of being much more computationally efficient, lending itself well to large-scale connected speech or real-time continuous speech recognition tasks.

Hidden Markov Models

A completely different approach to speech recognition is based on the assumption that speech can be modeled statistically. Indeed, it would be unreasonable to assume that all the information on a particular speech sound could be obtained from a single sample of that sound (which is essentially the approach taken in pattern-matching recognition schemes). By comparison, a statistically-based ASR system examines a large set of training speech data and generates a probabilistic model that characterizes the entire set. The resulting model of the speech unit is more powerful and general than a template.[29]

A particularly powerful stochastic process employs Hidden Markov Models (HMM) of each speech unit. In the HMM approach, speech is assumed to be a two-stage probabilistic process.[30] In the first part of the process, speech is modeled as a sequence of transitions through states. In the second part, the features of an observed speech unit are specified by a probability density function over the space of features. The Markov model is said to be hidden because one cannot directly observe which state the model is in. Only the features generated by the state are directly observable. The ideal HMM process models speech with the same variations that occur in human speech due to the effects of coarticulation, temporal distortion and other effects.[31]

During the recognition process, human generated speech is matched against an HMM by computing the probability that the HMM would have generated the same utterance or by finding the state sequence through the HMM that has the highest probability of producing the utterance.

---

28 Lee and Rabiner (1988)

29 Levinson and Roe (1990)

30 See Rabiner (1989) and Poritz (1988) for a clear, indepth discussion of HMM applications to speech recognition.

31 Peacocke and Graf (1990)

The literature indicates that there are some theoretical shortcomings of the HMM approach to ASR. Major among these is the false assumption that speech is a strictly Markovian process (in which the state transition probabilities are considered time-invariant). However, researchers have had considerable success in modifying the process to treat state transition probabilities as time-varying functions.[32] Another disadvantage of this process is that there is a significant probability that an observed utterance will be confused with an incorrect model. Again, research is on-going in an attempt to minimize this probability.[33]

The literature indicates that generally there is almost no difference in the recognition accuracy between the HMM and pattern-matching approach.[34] The major advantage of the HMM approach is that the process requires an order of magnitude less computation and storage capacity than the traditional pattern-matching scheme. However, the training phase for a HMM system is particularly long, due to the large number of samples of each word that are required to train the system. Conversely, the pattern-matching scheme requires only one sample of each word, consequently, the training phase is considerably shorter.

As was mentioned earlier in this paper, the two recognition schemes discussed above are the most highly developed and mature techniques employed in ASR technology to date. Variation of pattern-matching with DPG and HMM processes can be found in virtually all commercial ASR devices currently on the market, and are the only techniques that have demonstrated a level of accuracy compatible with the requirements of voice interactive cockpit technology.

VOCABULARY

The literature identifies two major characteristics of the system vocabulary that will directly affect the accuracy of the recognizer: vocabulary size and the phonetic similarities between different words in the vocabulary.

The larger the vocabulary, the greater the number of reference templates the recognizer will have to compare the input speech pattern to. Obviously, this will require a higher number of computations, and will slow down the recognition rate of the system. This has significant implications for real-time continuous speech recognition systems. It follows

---

32 Levinson and Roe (1990)

33 Moore, R.K. (1985)

34 Moore, R.K. (1985)

logically that as vocabulary size increases, the performance of the real-time system will degrade, causing delays in recognition rate that may be unacceptable for the required task. The system designer must therefore, carefully select a vocabulary that is large enough to meet the requirements of the task, but also sufficiently small enough for the computational capacity of the recognizer or host computer. It is worth noting that recognition systems based on HMM schemes are less affected by vocabulary size than are pattern-matching recognizers, since HMM processes are much more computationally efficient. With this in mind, it easy to understand why HMM schemes are finding widespread use in large vocabulary ASR systems.

A popular method employed in vocabulary development for connected and continuous speech recognition systems is to place syntatic constraints on the vocabulary itself.[35] That is to say, the vocabulary is structured so that only a certain subset of words are legal to be recognized at any one time, based on the first word of the phrase uttered. For example, returning to figure 10 above, starting at node (1), after the recognizer recognizes the word *I*, the only possible words that could follow at node (2) are *want, need, would,* and *will.* In this case, the recognizer is required to match the next utterance in the phrase against only four reference templates, as opposed to the forty-three templates that compose the entire vocabulary. Obviously, this scheme makes the recognition process much more computationally efficient, allowing for a larger working vocabulary. This syntax network type of vocabulary structure lends itself well to the distinct, rigidly structured control tasks characteristic of tactical aircraft cockpits.

Systems designers must also be careful in the selection of words or phrases that will make up the vocabulary. Obviously, recognizers will have a much more difficult time discriminating between phonetically similar words than widely dissimilar words. The designer must keep in mind the distortion introduced in the signal processing phase of the recognition process, as well as the expected S/N ratio of the incoming speech signal.

A good rule of thumb for vocabulary selection is to limit the size of the vocabulary to the degree that the vocabulary is still sufficiently large enough and flexible to meet the requirements of the task, but small enough to be managable by the recognizer. Word selection must take into account phonetic similarities, and every effort should be made to select words as dissimilar from each other as possible, within the constrains of task accomplishment. If the designer finds that the required vocabulary is too large or complicated for the signal processing and computational capabilities of the intended system,

---

[35] Casali, et al (1990),Levinson (1977), and Fu (1974).

he will be forced to finds methods to improve S/N ratio, refine signal processing techniques (to reduce distortion), and/or upgrade computational and memory capacity of the recognizer or host computer.

VOICE-INTERACTIVE COCKPIT RESEARCH

The application of speech recognition technology to the cockpit environment of an aircraft has been an area of intensive research over the past 20 years. The majority of this research has been in the form of laboratory and simulator-based studies, and has encompassed general aviation aircraft, military rotary-wing aircraft, civil transport and commercial aircraft, and military tactical jet aircraft and bombers. Analyses were made during these studies of what cockpit tasks were appropriate for the application of speech technology. The results of these analyses were very similar, and included tasks such as changing radio frequencies and waypoints, selection of display pages on MFD's, and retrieval of chart and procedural data.[36] Other studies relied on interviews and questionaires administered to experienced pilots in order to determine their preference for the application of speech technology to cockpit tasks.[37] These studies found that the preferred tasks were data entry and information retrieval. Interestingly, these studies also indicated that the more primary tasks, such as flight control activation, and arming and firing of weapons, were rarely identified as potential tasks for voice control.

A number of flight demonstration programs have also been conducted to evaluate the usefulness of voice-interactive cockpit systems and their ability to perform in the flight environment. These include the Advanced Fighter Technology Integration (AFTI)/F-16 Voice Interactive Avionics program,[38] a U.S. Navy sponsored F-18 program,[39] a French government sponsored Mirage aircraft project,[40] a British Buccaneer Aircraft

---

36 These conclusions were reached in studies concerned specifically with rotary-wing aircraft (Vidulich and Bortolussi, 1988), (Coler,1984), (Coler, et al, 1977); fighter cockpit simulators (Aretz, 1983); single-piloted aircraft operating under IFR conditions (North and Bergeron, 1984); and manned penetration bombers (North and Lea, 1982); and during more general studies reported by Gordon (1990), Anderson, et al (1985), LaPorte (1985), Moore, C.A. et al (1984), VanBronkhorst and Abraczinskas (1982), Montford and North (1980)and Hitchcock and Coler (1978).

37 Cotton, et al, (1983), Kersteen and Damos (1983)

38 Rosenhover (1987), Williamson (1987), Williamson and Curry (1984), and Werkowitz (1984).

39 Warner and Harris (1984).

40 Melocco (1984) and Tarnaud (1986)

Voice Recognition project,[41] the U.S. Army's Avionics Research and Development Activity (AVRADA) Voice Interactive Avionics program,[42] and the U.S. Army/NASA-Ames Research Center Helicopter Voice Recognition program.[43]

All of these programs have demonstrated remarkably similar results. The general conclusion is that state-of-the-art speech recognition systems are capable of limited operations, such as data input and information display functions, in the severe environment of a tactical aircraft cockpit. The studies also identified significant factors that will pose problems to any increase in the use of voice interactive systems to other cockpit tasks. These factors can be broken down into three major areas: the flight environment, the aircraft equipment, and human factors.

FLIGHT ENVIRONMENT

It is not uncommon for new problems to arise when technology is moved from the controlled environment of the laboratory to the field. This has particularly been the case with the application of speech recognition technology to the severe environment of the aircraft cockpit. The literature identifies three primary characteristics of the flight environment that will impact upon the performance of ASR systems: noise, acceleration, and vibration. Not only will these 'stressors' physically affect the performance of the recognizer itself; studies have shown that their effect on the pilot's ability to produce speech samples that are reasonably consistent during any and all phases of flight is even more pronounced and pose a much more difficult problem.[44] The physical effect of aircraft acceleration and vibration can be minimized by proper packaging and flight-worthy design of the ASR system. However, the variations in speech production caused by acceleration and vibration are more difficult to minimize, and recognition algorithms must necessarily be more flexible to contend with these variations. Noise can pose an even greater problem. Noise levels as high as 108 dB have been recorded in the AV-8B cockpit. This high of a noise level can virtually obscure word boundaries, making connected and continuous speech recognition practically impossible. However, it has been demonstrated that algorithms can be developed to handle up to 112dB of noise, matched in spectra to that measured in an F-16 cockpit in flight.[45] Noise-cancelling microphones that filter out

---

41 Smith and Bowen (1986).

42 Pondaco (1989), Westerhoff and Reed (1985).

43 Simpson (1991).

44 Moore, T.J. (1989).

45 Williamson and Curry (1984).

background noise and improve S/N have also been investigated and are constantly being improved and refined.[46]

The effect of noise on pilot speech production is one of the more interesting and difficult problems for cockpit ASR integration. It has been discovered that in the presence of high ambient noise levels, speakers tend to increase the volume of their speech and to increase their fundamental frequency. This effect, called the 'Lombard Effect', is well documented in the literature.[47] In addition to the increases in volume and frequency of speech, there are also effects on the formants; the distribution of energy within the speech spectrum shows an increase in the high frequency third formant (F3) component, and the vowel space between F1 and F2 gets smaller. In terms of the effects on speech recognizers, it has been demonstrated that the effect of noise at the speaker's ear results in a greater degradation in the performance of an ASR device than does the presence of noise at the speaker's microphone.[48] The implication of this study being that changes in speech production due to high cockpit noise levels are sufficient to affect the performance of existing pattern-matching speech recognizers. However, recent flight evaluations of noise-canceling earphones and flight helmets have shown significant promise of minimizing the Lombard effect.[49]

The effects on speech production of the high levels of acceleration typically experienced in tactical jet aircraft have also investigated. A study of two male speakers wearing oxygen masks and chest mounted breathing regulators was conducted at normal 1 g conditions and at sustained accelerations up to +6 g to obtain information about the effects of acceleration on the acoustic-phonetic structure of speech.[50] Increases were found in the fundamental frequency for both speakers. Similar to the effect of noise, acceleration also caused the vowel space defined by F1 and F2 to decrease. These results seem to suggest that acceleration will have much the same effect as noise on the performance of current speech recognizers.

Studies of airframe vibration effects have seen widely varying results. Laboratory studies as well as data collected during flight test evaluations indicate that current helicopter vibration environments doe not substantially affect the performance of selected ASR

---

46 Endres (1991), Edgerton (1986), Singer (1981) and Lu, et al (1980)

47 Bond, et al (1986), Pisoni, et al (1985)

48 Rajasekaren, et al (1986), Landell, et al (1986), Rollins and Wiesen (1983).

49 Simpson (1991), Rajasekaren and Doddington (1985)

50 Bond and Anderson (1987)

systems.[51] However, a study conducted as a ground evaluation of the ASR system installed in the AFTI/F-16 yielded quite different results.[52] During this evaluation, speakers wore oxygen masks and were subject to four different levels of vibration: zero vibration, and low, medium and high vibration. The vibration modes were set up to emulate as closely as possible the airframe buffet that would be experienced on a low-level, high speed flight. Not surprisingly, a modulation or shakiness was imposed on the voice, due to the whole body vibration. More significant however, was that once again, as was the case with noise and acceleration effects, the fundamental frequency increased, and the vowel space became more compact.

In summary then, it appears that cockpit noise, aircraft acceleration and airframe vibration will have some effect on the performance of cockpit ASR systems. The extent to which these stressors will degrade performance seems to be very much 'aircraft dependent'. It is more likely that ASR performance will be more affected in the tactical jet aircraft environment than say, the civilian transport or airliner environment. It follows then, that cockpit ASR system integration must make a strict and through assessment of the expected operating environment and tailor the system to minimize the effects of that environment.

AIRCRAFT EQUIPMENT

The equipment with which the ASR system will have to interface is an important consideration for system integration. For all applications, the microphone is an important factor in the overall performance of the recognizer. Ideally, it would be best to chose the highest quality and best performing microphone suitable to a particular application. However, it is more likely that the ASR system will have to interface with the existing equipment. For tactical jet aircraft applications, this means the oxygen mask/M101 microphone combination.

The effects of wearing an oxygen mask/M101 microphone combination have been studied and two significant problems have been identified.[53] The first and most dramatic effect was that the vowel space for all nine vowels studied was compressed in the F1 dimension. An unrelated study showed that when jaw movement is restricted, such as is the case when wearing an oxygen mask, the vocal effort is reduced and other articulatory

---

51 See Dennison (1987) and Cruise, et al (1986) for laboratory studies; Malkin and Dennison for results of in-flight evaluation.

52 Moore and Bond (1987)

53 Moore and Bond (1987), Malkin (1986), and Singer (1981)

compensations must be made, perhaps resulting in the compression of the overall vowel space and tighter clustering.[54]

The second problem posed by the oxygen mask is the noise associated with the valve motion when the pilot breathes in and out. Fortunately, current pattern matching algorithms can model this noise and reject it during the recognition process. However, this means that the ASR system would have to be trained with the pilot wearing the oxygen mask/microphone combination he expects to wear in the aircraft.

Another important factor for consideration in ASR/aircraft equipment interface is the intercom system onboard the aircraft. As was discovered during the AFTI/F-16 program, current aircraft intercom systems are not designed to meet the requirements of ASR systems.[55] In the case of the F-16, it was necessary to install a separate amplifier in parallel with the existing intercom system.

In summary, it is more likely that during ASR system integration, the ASR system will have to be tailored to meet the interface requirements of the aircraft equipment, rather than the aircraft equipment being modified to meet the interface requirements of the ASR system. The degree to which existing aircraft equipment will impact recognizer performance will depend on many factors, and is obviously very much aircraft dependent. Once again, it is imperative for the cockpit ASR system designers to accurately assess all factors of a particular aircraft's equipment that may impact on ASR system performance, and tailor the system to minimize the effects of those factors.

HUMAN FACTORS

An area of great concern and extensive research is the human factors issues involved in the design and integration of a voice-interactive system into the cockpit. Simply stated, it is difficult to predict, let alone model, human behavior under a variety of flight conditions and stress levels. To date, research seems to have uncovered more questions than answers.

One of the more challenging questions in systems integration is the display format. More specifically, pilot/cockpit interface design issues include: determination of what modes of feedback are most appropriate; the merits of visual versus audio displays; and stand-alone displays versus integrated displays.

---

54 Schulman (1985)

55 Rosenhover (1987)

Research conducted at the U.S. Army Research and Technology Laboratories uncovered two major problems in integrated displays which were not found in stand-alone visual or speech displays.[56] These problems are display priority and temporal veridicality.

Display priority problems are a result of a fundamental limitation of voice displays. Voice displays can only present one item of information at a time, whereas visual displays can usually present a great deal of information simultaneously. Therefore, prioritization of information is much more important in voice displays than in visual displays. Temporal veridicality is a result of the time lag involved with voice displays. Visual displays present information instantaneously; however, there is an unavoidable delay in audio feedback of information. In integrated displays, a conflict can quickly arise when a visual display presents something different from the associated voice display. Obviously, this can lead to pilot confusion, frustration, and lack of confidence in the voice system.

Apart from the pilot/aircraft interface issues, other human factors that could affect the performance of the ASR system are pilot emotional stress and speaker variability.[57] It has been well documented in a number of studies that an individual's emotional state often results in changes in the acoustic characteristics of their speech. The most consistent and significant changes are an increase in the frequency and variability of the fundamental frequency.[58] Lincoln Laboratories have reported some success in training ASR systems using training tapes of a set of speech samples of an operator speaking in a variety of styles (fast, slow, shouting, etc.), and generating templates by an HMM averaging technique. The templates appear to encompass the variability seen in speech produced under emotionally stressful conditions.[59]

It is well documented in the literature that not all individuals will have equal success when working with ASR technology. The AFTI/F-16 flight evaluation reported that ASR performance varied significantly from pilot to pilot.[60] Studies indicate that for 80-90% of the population satisfactory results can be achieved for ASR, while the remaining 10-20% will always perform poorly with the current technology.[61] However, other studies have indicated that with proper feedback and training on the system, a speaker's performance

---

56 Voorhees and Bucher (1985), Moore, C.A. and Ruth (1984), Simpson, et al (1982)

57 Moore, T.J. (1989)

58 A complete review of the literature dealing with this topic can be found in Williams and Stevens (1981)

59 Paul, et al. (1986)

60 Williamson (1987)

61 Doddington (1986)

with an ASR system can be improved in a reasonable amount of time.[62] The implication here being that training and experience with the ASR system is an important factor in the success and user acceptance of the technology.

The conclusions to be drawn from the literature on the subject of human factors in voice interactive cockpit research is that any design effort must include a thorough assessment of the intended use of the system, the tasks or missions to be accomplished, the pilot population that will use the system, and the flight environment to which the system will be subjected, prior to actual design and integration into the aircraft. Resolving the human factors issues of any particular VIS design will require the collective effort of the systems engineers, linguists, aviation physiologists and cognitive psychologists.

SUMMARY

The literature indicates that for the foreseeable future, pattern-matching, speaker-dependent, small vocabulary, connected and continuous speech recognition systems are the most attractive systems for cockpit applications. HMM-based systems, although for years considered more computationally efficient, are losing the computational advantage over pattern-matching systems due to the tremendous increase in computing power of microprocessors. The most attractive advantage of pattern-matching systems is the relatively simple system training process. Whereas HMM systems require several samples of each speech token, pattern-matching systems require only one. Obviously, the more simple and straight forward the training process, the higher the user acceptance rate will be.

The limited vocabulary of current pattern-matching continuous speech recognizers does not appear to be an obstacle for tactical aircraft integration. It is of greater benefit to the pilot to keep the vocabulary as small as possible, so that there are fewer commands to remember. Through proper syntatic network design, the vocabulary may be made small, yet flexible and natural enough for the pilot to use.

Current military microphones and headsets are not adequate to meet the needs of state-of-the-art ASR systems. While it is reasonable to assume that the next generation fighters will incorporate noise-cancelling headsets and microphones, it would be unrealistic to require the military to modify or redesign and replace existing equipment. Such an option would be too expensive. ASR integration will therefore, have to take the aircraft equipment

---

62 Zoltan-Ford (1984)

factors into account and modify the signal processing and noise rejection algorithms to contend with the effects of these factors.

Human factors issues will continue to be an area of extensive research. While there are common trends in all the studies conducted to date, each application of ASR technology to a particular aircraft will define a unique set of problems.

Despite the issues mentioned above, ASR technology is at the point where it is ready to be integrated into tactical jet aircraft. All of the flight test evaluations mentioned above have demonstrated that, given a limited vocabulary and appropriate syntactic constraints, recognition rates of greater than 90% are achievable. They have also shown that the use of speech technology in the cockpit results in a decrease in workload. These results suggest that it is feasible to consider the integration of speech technology into the AV-8B aircraft.

IV. METHODOLOGY

INTRODUCTION

The test aircraft selected for the Harrier Voice Interactive Technology Program was a TAV-8B two-seat testbed aircraft, located at and operated by the Naval Air Test Center (NATC), Patuxent River, Maryland. The goal of the program was to explore the feasibility of integrating current VIS technology into operational Harrier aircraft within two years, and evaluate the potential of the system to reduce pilot workload.

Constraints placed on the program by the program sponsor, NAVAIRSYSCOM, restricted the modification of any existing cockpit hardware, including the pilot's microphone and headset. The philosophy for imposing this constraint was to determine if VIS could be integrated into operational aircraft with minimal modifications to existing equipment. However, additional cockpit hardware could be added for the purposes of the evaluation. Changes to the Mission Computer (MC) and Display Computer (DP) software were allowed, and of course, allowances were made for the installation of the VIS hardware in the equipment bay of the aircraft. An additional limitation placed on the integration effort was that the VIS vocabulary had to emulate exactly, the pushbutton methodology of the cockpit controls and displays. The reasoning for specifying the vocabulary in this manner follows from the assumption that Harrier pilots were already familiar with the switchology of the cockpit controls and displays, and therefore training time with the new system would be greatly reduced.

The following paragraphs describe the test aircraft, test equipment, vocabulary, system integration, and test methodology of the evaluation program.

DESCRIPTION OF TEST AIRCRAFT

The TAV-8B, illustrated in figure 11, was a two-seat testbed aircraft modified from the production AV-8B aircraft. The initial cockpit was moved forward and an identical cockpit was installed above and behind the original. The rear cockpit contained identical displays, HUD, UFC, and control stick and throttle placement. Maximum commonality between the TAV-8B and AV-8B aircraft was maintained for test purposes.

Aircraft avionics were interfaced through use of two Military Standard (MIL-STD) 1553B multiplex data buses (mux bus) which contained redundant bus lines. Bus control was maintained by an AYK-14 /XN-6 mission computer.

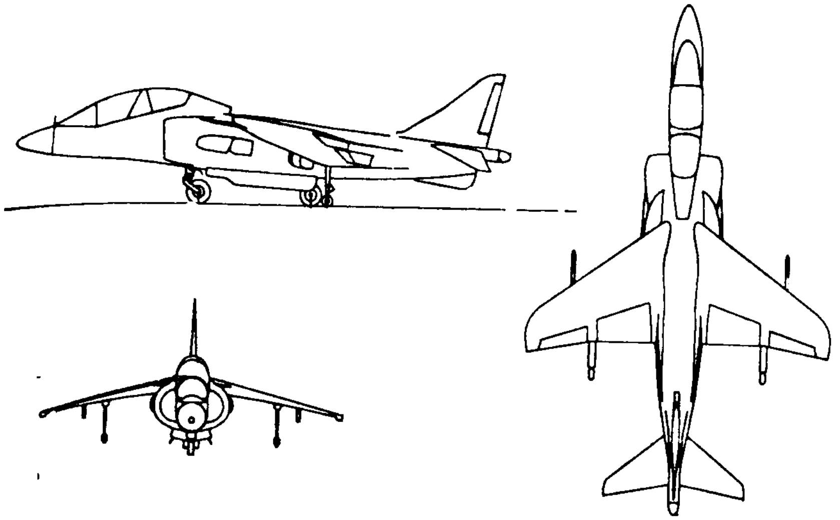


Figure 11
TAV-8B TESTBED AIRCRAFT

DESCRIPTION OF TEST EQUIPMENT

The VIS used for this evaluation consisted of a VRS-1280 VRS, produced by ITT Defense Electronics Corporation, integrated into a SANDAC V host computer, which was manufactured by the Sandia Corporation. Speech synthesis functions were performed by an ITT Continuous Variable Slope Deltamod (CVSD) text-to-speech type synthesizer, which was also installed as part of the VRS-1280, and utilized the host computer's memory for storage of speech text.

The VRS-1280 was designed as a speaker dependent continuous speech pattern-matching recognizer. The system operated on a finite state grammar and recognition results were output on a node-by-node basis as the unknown utterance was being spoken. The recognizer was designed to achieve real-time operation with up to 800 words in the syntax directed vocabulary. The VRS-1280 was packaged on a single PC board. Physical dimensions and systems capabilities are listed in table I.

A block diagram of the system architecture is shown in figure 12. Audio input from the pilot's microphone was amplified then sampled and digitized at an 8 KHz rate by a digital Coder/Decoder (CODEC) chip. The signal was then sent to a Digital Signal Processor (DSP) which performed a bandpass filter (BPF) operation, calculating the relative energy in fourteen frequency bandwidths every 10 msec. The BPF coefficients were subsequently

Table I
VRS-1280 DIMENSIONS AND CAPABILITIES

Recognition Vocabulary	500 words 2000 operational templates
Recognition Throughput	800 words
Syntax Nodes	1024
Synthesis Method	16 Kbps CVSD
Length	7.00 inches
Width	0.57 inches
Height	6.25 inches
Weight	1.0 lb
Power Requirement	5 watts

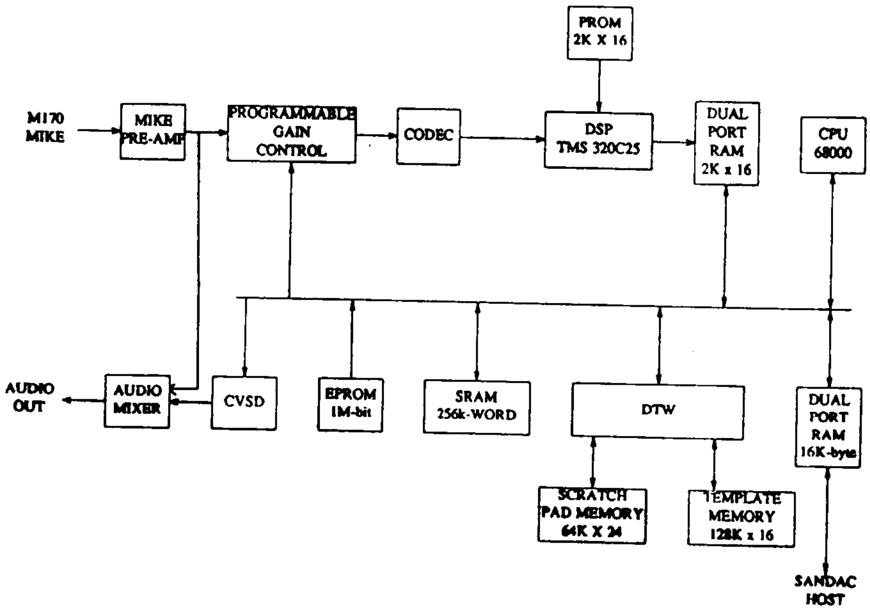


Figure 12
VRS-1280 SYSTEM ARCHITECTURE

processed by a 68000 Central Processing Unit (CPU) which gain normalized the coefficients and reduced the data to eight mel-cepstral coefficients every 20 msec. These eight parameters were the basic elements upon which speech recognition decisions were based.

During the recognition process, the frame data was sent to a DTW chip which performed the time warp algorithm and compared the unknown frame of data with the template frame data and output a distance score. The best scoring templates were identified and stored in a list of the best scoring phrase options. The syntax-directed options determined which words or templates would be compared to the next frame of unknown data. Recognition results were output on a node-by-node basis.

It is important to note at this point that the VRS-1280 was a syntax-directed 'phrase' recognizer vice an 'isolated word' recognizer. The recognition process, once activated, continuously evaluated all possible paths through the application's vocabulary. In this context a path was defined as any possible sequence from a particular node (word) to any one of the possible endpoint nodes (endpoint words). Each path represented a sequence of words which made up a phrase. The path which contained the sequence of best scoring word templates was recognized as the phrase uttered by the pilot.

VOCABULARY DEVELOPMENT

The development of the VIS vocabulary was greatly simplified by the constraint placed on the program that the vocabulary had to exactly emulate the manual pushbutton switchology of the existing AV-8B cockpit workstation. Since cockpit switchology already followed a syntax-directed or network type of methodology, the vocabulary structure was in effect already defined.

Functions selected for voice activation by the test team included communications, ECM, navigation, weapons programming, target data entry, and target timing options. The decision was made not to include any functions that could be activated by the HOTAS controls since these controls were already immediately accessible to the pilot. Additionally, weapons arming and release functions were not included for obvious safety considerations.

Three new aircraft status functions were incorporated into the system: altitude, fuel state, and combat. The status functions were designed to respond to a pilot query with synthesized speech feedback. The combat status function in particular was designed to check on proper status of all required weapons and systems prior to target attack. The VIS was designed to reply *combat* if the systems were up and ready, and *master arm*, *air-to-*

ground, or *weapon* depending on if, respectively the master arm switch or air-to-ground mode were not selected, or the weapon was not selected or programmed properly.

The limitation of not modifying existing cockpit hardware forced the implementation of a keyword to activate the VIS before speaking a command. The keyword had to be phonetically distinct in order for the VIS to have a high level of confidence that the word spoken was actually the keyword. In effect the system was listening all the time, however it would not begin the recognition process until the keyword was spoken and recognized. The keyword selected for this evaluation was *VIC-ON*, which is an abbreviation for Voice Interactive Communications-On.

The selected VIS vocabulary, depicted in figure 13, contained 76 unique words arranged in a syntax-directed finite state grammar network. Each page represented a node in the network. The word *disengage* was included in the vocabulary to deactivate the system at any time. An example of how the vocabulary works, a valid VIS command to select a radio frequency on the radio #1 would be: *VIC-ON COMM-1 MANUAL THREE ONE FOUR POINT FIVE ENTER*. Likewise, the proper command for selecting a navigation waypoint by voice would be: *VIC-ON WAYPOINT SEVEN*. A detailed description of all VIS systems functions, a complete list of valid phrases and commands, and illustrations of all possible paths through the vocabulary is contained in appendix A.

TEMPLATE DEVELOPMENT

A template training system was developed to train the speaker dependent vocabulary templates and create a synthesis library. The system consisted of a VRS-1280 PC board installed in a COMPAQ 286 portable personal computer. A software-based training tutorial developed by AVRADA for an earlier helicopter VIS evaluation was modified for the Harrier-specific vocabulary. The tutorial guided the pilot through the vocabulary by prompting him to say each word in the vocabulary, both as an isolated word and as part of a phrase. In order to minimize coarticulation effects, each word was included as the first, middle and last word in three separate phrases. The system then had four samples of each word. These samples were averaged together to create a single template of each word. During signal processing of the pilot's speech sample the system measured the energy content of the signal spectrum and rejected the sample if the energy content was too low to define a distinct speech pattern.

Once the system accepted a sample of each word, the tutorial guided the pilot through a recognition test of his voice templates by prompting him to say each word in the vocabulary and matching the utterance with one of his newly created templates. A distance score was

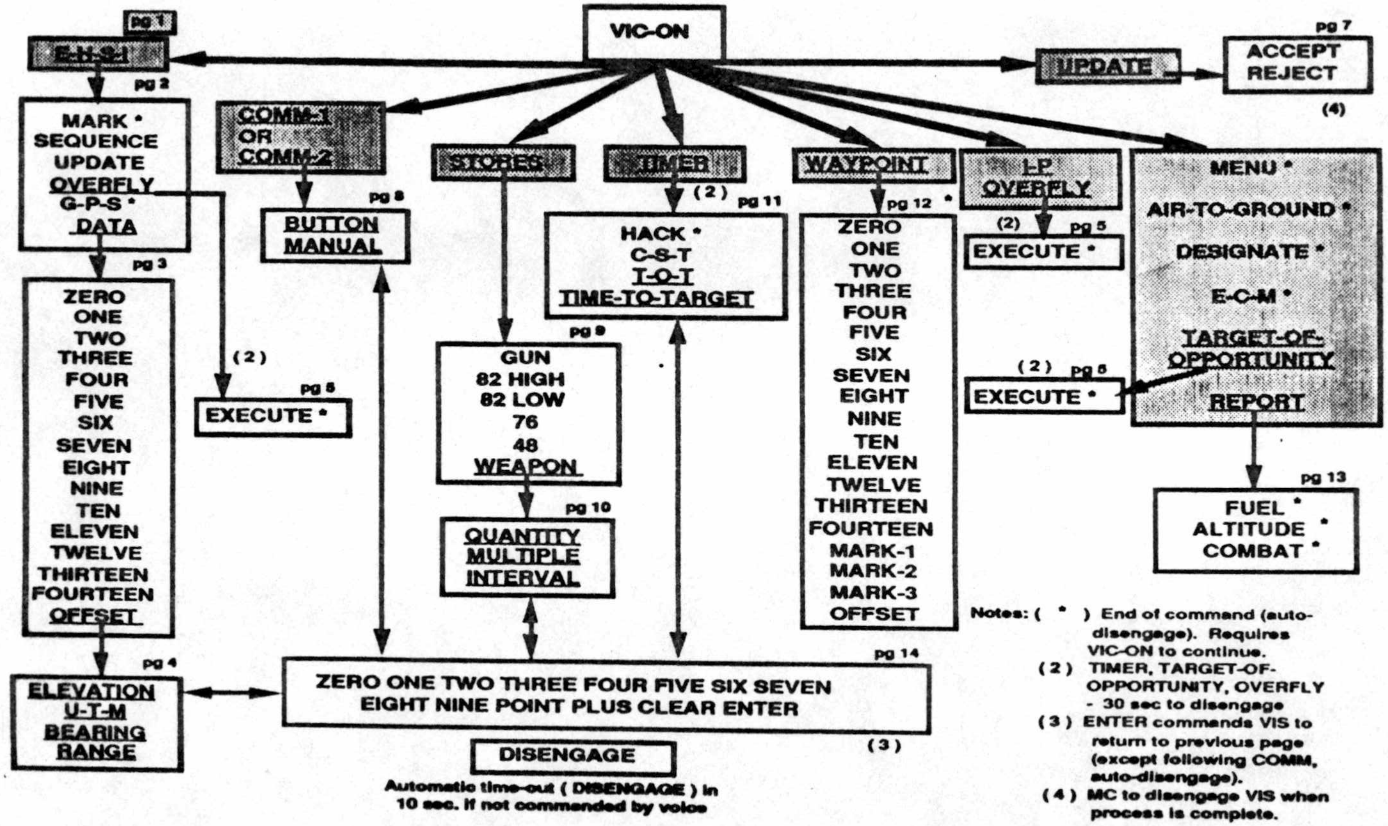


Figure 13
VOICE INTERACTIVE SYSTEM VOCABULARY

output with each match. If the score was too low the tutorial prompted the pilot to give another four samples of the word. The system then: 1) averaged the samples and stored this new template, 2) prompted the pilot to once again say the word, 3) matched the utterance to the new template and 4) output a new score. This process continued until an acceptable score was obtained. Once this process was complete, the recognition test resumed through the vocabulary until a complete set of high scoring templates for each word in the vocabulary was obtained.

Because the VIS was designed to be activated by a keyword and would in effect listen all the time for that keyword, a set of noise rejection templates had to be developed to match and then reject invalid speech and non-speech sounds. Without these generic templates, the VIS would attempt to match any sound it heard with a voice template, resulting in false activation of the system. Twenty generic noise templates containing random phonemes were developed and included as part of the template library.

After the vocabulary had been developed and verified, test pilots evaluated the recognition performance of the VRS-1280 by speaking valid command phrases while wearing their helmets and oxygen masks. It was soon discovered that the noise created by the continuous stream of oxygen flowing over the microphone, coupled with the motion of the valve of the oxygen mask opening and closing, significantly degraded recognition of the valid vocabulary commands.

To alleviate this problem, a set of generic breath templates were generated to model the noises associated with the oxygen mask. First a set of vocabulary templates were made with the pilot wearing his oxygen mask, but in a quiet environment with the oxygen regulator disconnected from the mask. Then, another set of vocabulary templates were made with the regulator reattached to the mask. The regulator was attached to a bottle of compressed air to simulate pressure breathing in the flight environment. To simulate the acoustic environment of the Harrier cockpit, the entire process was conducted in a noise chamber which played recorded tapes of cockpit noise in various phases of flight. Sounds levels varied from 105 to 110 dB. The final step in the process was to subtract the quiet vocabulary template patterns from the noisy templates, leaving only the signal patterns of the breathing noises intact. Once these breath templates were incorporated into the VIS, recognition accuracy increased to 100% on the template training system.

The final step in the template development process was to load the voice templates onto a 3 1/2 inch computer disk for transfer to the aircraft VRS-1280. Since the system was speaker dependent, each pilot was assigned a disk which contained his voice templates of the vocabulary. Noise rejection and breath templates were loaded onto a separate disk for loading into the aircraft system.

Speech patterns were also created for speech synthesis functions. The synthesis library was small for this evaluation and consisted of the words: *ready*, *error*, *disengage*, *combat*, *speak*, *quiet*, *calibration*, *complete*, *hundred*, *thousand*, and the numbers *zero* through *nine*. The patterns were created by a female speaker who spoke each word into the template training system. The system digitized the speech pattern for each word and output the processed pattern onto a 3 1/2 inch computer disk for loading into the host computer's memory.

AIRCRAFT INTEGRATION

The integration of VIS into the testbed aircraft required modification of the MC software, installation of the SANDAC V computer, and the addition of a VIS control box to the forward cockpit.

As shown in figure 14, the design approach taken to integrating VIS into the avionics system of the aircraft was to install the VIS hardware in parallel with the cockpit controls and displays. The VIS sent an identical signal to the MC that the associated cockpit switch or display pushbutton would have sent, had it been manually actuated. It was indistinuishable to the MC whether the signal came from VIS or a manual cockpit control. This approach had the advantage of not having to modify the MC or mux bus architecture.

The block diagram of the VIS integration is shown in figure 15. For ease of integration, only the forward cockpit audio system was wired into the VIS. Audio was transmitted

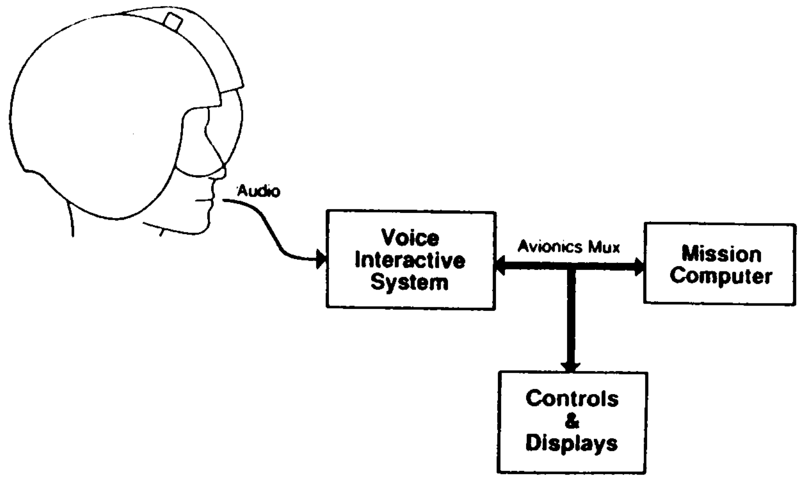


Figure 14
VOICE INTERACTIVE SYSTEM OPERATION

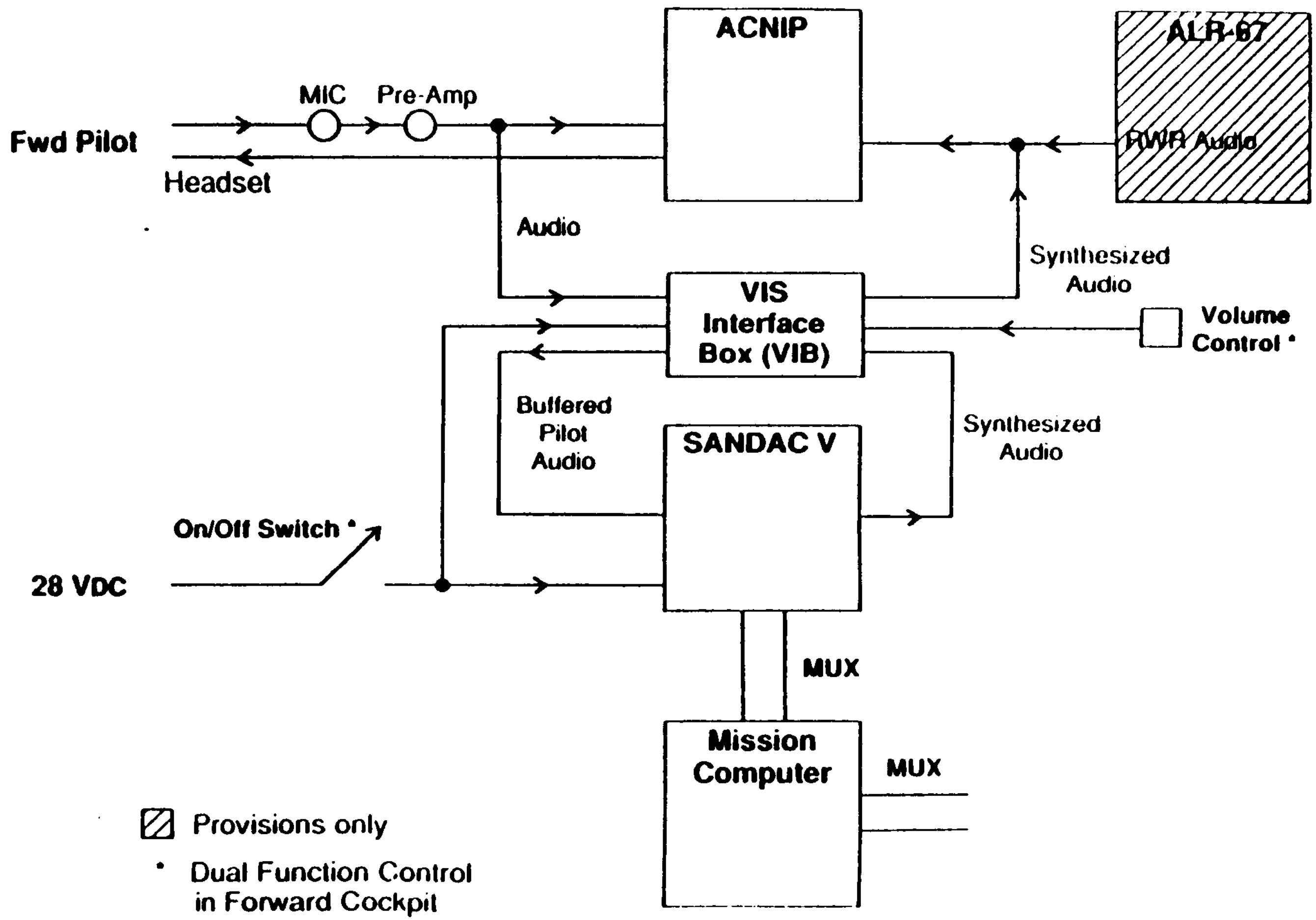


Figure 15
VOICE INTERACTIVE SYSTEM BLOCK DIAGRAM

from the pilot's microphone through a pre-amplifier to the VIS Interface Box (VIB), which was embedded in the SANDAC V. The VIB contained the VRS-1280 recognition and speech synthesis cards. The audio was also fed to the Auxiliary Communications Navigation Identification Panel (ACNIP), which provided sidetone feedback to the pilot's headset as well as Radar Warning Receiver (RWR) tones (based on signals from the ALR-67). Once a valid command was recognized by the VIB, it sent a coded signal to the SANDAC V computer. The SANDAC acted upon this signal, sending an identifiable signal to the MC. The MC then acted on that signal, sending a signal through the mux to the DP to change a display or input data to the avionics and/or weapons systems.

The SANDAC V was designed to performed a variety of functions. It acted as the interface between the VIB and the MC, and sent appropriate signals which would be recognizable to the MC software. It was also designed to queried the MC on various systems status when prompted by the pilot's verbal request, sending the appropriate signals back to the VIB for speech synthesis operations.

Other SANDAC V design functions included commanding the VIB to send synthesized audio to the pilot's headset of VIS status. The word *ready* was heard by the pilot when he said *VIC-ON* and the system was ready to begin the recognition process. Similarly, the word *disengage* would be heard when the pilot commanded system disengagement by saying the word *disengage*, or when the recognition process was interrupted before reaching an endpoint word. Also, the SANDAC's software was designed to send an error signal to the pilot when a words or phrases were spoken out of sequence. In this case, the pilot would hear the synthesized phrase *error,error* over the headset.

The SANDAC V was designed to terminate the recognition process and reset the VIS to await another verbal command when any one of four conditions were met: 1) the VRS-1280 reached an endpoint word of a valid command, 2) the word *disengage* was recognized, 3) the SANDAC detected an invalid command, or 4) when the system 'timed out'. In all but the first case, the word *disengage* was signaled to the pilot upon termination of the recognition process.

The time out function was incorporated into the VIS to automatically disengage the system after 10 seconds had elapsed from the time that the last recognized word had been uttered by the pilot. This feature was included so that in the event that the pilot stopped speaking before uttering the entire sequence of words that made up the command, the system would not be actively listening to extraneous non-speech sounds and attempt to come up with a match in the vocabulary, which would have resulted in an insertion error or generation of a false command.

An exception to the 10 second time out function was made for the waypoint overfly update, Initial Point (IP) overfly update, and Target Of Opportunity (TOO) functions. These functions were used to mark the position of a ground reference with the Inertial Navigation System (INS). To manually execute these functions, the pilot maneuvered the aircraft to fly over the ground reference or target, then selected the overfly pushbutton on the Options Display Unit, or in the case of the TOO, the TOO pushbutton on the UFC. Upon actuation of one of these two pushbuttons the INS stored the target's latitude and longitude in the MC. The MC displayed time, distance, heading and position information on the target to the pilot on the HUD and MPD.

For voice actuation of these functions, it was determined that 10 seconds would not be enough time for the pilot to maneuver the aircraft over the ground reference before the selected function timed out. Therefore, the time-out duration for these three functions was increased to 30 seconds. This was determined during the integration phase of the project to be adequate time for the pilot to 1) call up the function, i.e., say *waypoint overfly*, and then 2) maneuver his aircraft over the point and say *execute* to mark the spot with the INS.

Modifications to the existing MC software included the installation of mux bus addresses to enable the SANDAC V to ascertain from the MC the status of aircraft altitude, fuel state, weapons system go/no-go status, and master armament switch position (this data was required for the VIS status response functions). Also included in the software revision was the addition of a Voice Calibration (VCAL) function, which was selectable on the MPD. The system was designed so that when the pilot selected VCAL, the VIS would sample the ambient noise level of the cockpit at the pilot's microphone. The system would then prompt the pilot to speak a canned phrase. The system would then adjust the gain of the audio system for the optimum S/N.

Hardware additions to the aircraft included the installation of a rotary VIS volume control knob, mounted on a 6x6x8 inch control box, located on the rear portion of the right console of the front cockpit. The box occupied previously unused space and was placed so as not interfere with the execution of normal cockpit tasks during flight.

The only other hardware addition to the aircraft was the installation of the SANDAC V computer and its associated wiring. The computer, which measured 7.0x10.7x6.5 inches and weighed 22 pounds, was placed in the rear equipment bay of the aircraft, and did not significantly effect the weight or center of gravity of the aircraft. Special wiring for the SANDAC interfaces between the cockpit and the MC followed the normal wire bundle pathways of the aircraft and did not interfere with the operation or maintenance of aircraft equipment.

TEST METHODOLOGY

The test methodology included subjective and objective measures of VIS performance and pilot workload reduction. Two test pilots with in-depth operational experience flying the Harrier in the tactical flight environment were selected to participate in the project. A total of 12 evaluation flights were planned, divided equally between the two pilots. The objectives of the evaluation program were: 1) determine the recognition accuracy and system performance of the VIS under a variety of flight conditions, 2) subjectively evaluate the degree of workload reduction in the cockpit while using voice-actuated controls and audio displays, 3) identify which cockpit workstation functions were best suited for voice control, and 4) evaluate the potential of VIS technology to further reduce cockpit workload and increase pilot flexibility,

The flight test plan consisted of six flights per pilot, flown from the forward cockpit of the aircraft The second pilot occupied the rear cockpit, acting as a safety pilot and observer. All flights were flown in daylight Visual Meteorological Conditions (VMC). The first three flights consisted of an operational checkout of the system in various modes of flight. During this phase of testing the pilot was required to say each of the 59 different valid commands during the takeoff, cruise, high-g maneuvering, and approach and landing phases of flight. A detailed description of test work and procedures can be found in appendix B.

The next three flights consisted of a subjective evaluation of the VIS in the tactical flight environment, employing the aircraft in a Close Air Support (CAS) scenario. The CAS mission was considered the highest workload mission profile flown by operational Harrier pilots. In this scenario the pilot manned up the aircraft and stood in an alert status until directed by the Direct Air Support Center (DASC) to launch on a mission. Once airborne, the DASC directed the pilot to proceed to a specified Control Point (CP) and contact the Tactical Air Controller (Airborne) (TAC(A)). After a short loiter over the CP, the TAC(A) directed the pilot to contact the Forward Air Controller (Airborne) (FAC(A)) for a 9-line CAS mission brief. After confirming the 9-line brief, the pilot entered: the target location's latitude, longitude and UTM coordinates, weapons delivery data for the specific target, and time the bombs were assigned to be on target. Additionally, the pilot entered the Initial Point (IP) location and the course and distance from the IP to the target.

While the pilot was entering the required data, he also had to leave the CP at the correct time to arrive at the IP and execute the IP to target run to place the bombs on target at the assigned time, within a 20 second window. The ingress to the target was flown at 200 feet

AGL at 480 KIAS. After simulated release of the ordnance on target the pilot was required to egress from the target area at 200 feet AGL to a Rendevous Point (RP), where he was directed to contact the TAC(A) and await assignment of another CP and CAS mission brief. A detailed description of the CAS scenario is contained in appendix C.

During the CAS evaluation, the pilot first flew the mission profile using only manual methods of data entry and cockpit management. Then the pilot flew the same profile using the VIS to aid in management of cockpit tasks during the preparation and execution of the mission. This method was intended to allow a clear and immediate comparison of the usefulness of VIS technology and help identify the best cockpit functions for VIS application.

Data collection devices consisted of a video and audio cockpit recorder, and pilot data cards for pilot comments. The cockpit recorder recorded the HUD and MPD video, pilot voice commands and queries, and VIS/aircraft response. Cockpit tapes were played back immediately following the flight for data reduction. The test team recorded the number of commands spoken by the pilot, and the number of correctly and incorrectly recognized commands, insertion errors,false starts and ignored commands. Pilot comments were also recorded on the cockpit recorder and transcribed during the data reduction process. Pilot interviews were conducted immediately following the evaluation flight to record their initial impressions on the system and the usefulness of the technology.

V. RESULTS

GENERAL

The flight evaluations revealed some interesting results. Contrary to what the test team expected, the vocabulary proved to be cumbersome and difficult to use. Especially under high workload phases of flight, pilots found it difficult to remember the proper sequence of words that made up each command. The voice activated switch command, VIC-ON, was not recognized during a large number of voice commands, leading to a high number of ignored commands. Forty-nine percent of all commands were ignored by the VIS.

This high number of ignored commands had a adverse impact on the recognition accuracy of the system. However, if recognition accuracy is considered for only those instances in which the system recognized the command VIC-ON and initiated the recognition process, the recognition accuracy was 95.7%. The VIS had the most difficulty recognizing long digit strings (7 or 8 digits), and the highest percentage of mis-recognized commands involved long strings of numbers, such as seven digit UTM coordinates.

False activations of the system were rare. During 12 flights totaling 16.8 flight hours, 11 incidences of false system activation were reported. Of these, seven occurred during one flight. A faulty voice calibration was suspected by the test team and a recalibration in-flight resulted in no false activations for the remaining 41 minutes of the flight. Insertion errors were likewise rare and were limited to insertion of digits in long digit strings.

The evaluation pilots reported difficulty using the timer and the navigation system update functions. These are 'time-critical' functions. In the case of the timer function, the pilot was required to say *hack* when the FAC(A) gave the time hack over the radio. Since the timing of the weapons delivery on target depended solely on the time hack, the mis-recognition of the command *hack* resulted in significant delay in initiating the timing functions, and caused the pilot to be late to the target . Interestingly, the word *hack* was only mis-recognized by the VIS during the high workload CAS scenario and not at all during the relatively low workload operational checkout flights. Out of 24 timer commands given during the CAS flights, nine were either mis-recognized or ignored.

Likewise, the time critical nature of the navigation system overfly updates caused the pilots some difficulty. These functions incorporated a 30 second delay feature, intended to allow the pilot to verbally call up the function while maneuvering to fly over top of the target and then, when immediately overhead, say the command *execute* to mark the location. The pilots reported difficulty in timing the call up of the function so as to arrive over top of the ground reference before the 30 seconds allotted had timed out. In four instances, the function timed out when the pilot was within five seconds of overflying the ground reference, and was forced to quickly call the function back up and execute the overfly update. In these instances the pilots found that before they could recall the function and execute it, they had already flown over the ground reference and were required to either fly back over the point or reject the update and continue on with the mission profile.

Functions that were found to be very helpful and demonstrated recognition accuracies approaching 100% were the communications, waypoint selection and status report functions. The waypoint selection and status report functions in particular were singled out for favorable comment because they provided the pilot with a capability that did not previously exist. The waypoint selection function allowed the pilot to select a waypoint out of sequence, as opposed to the manual method of scrolling through waypoints sequentially until the desired waypoint is found. The status report functions allowed the pilot to ascertain aircraft altitude, fuel state, and readiness for ground attack without having to look inside the cockpit. The evaluation pilots found this gave them much more flexibility in cockpit management tasks during high workload phases of flight.

SPECIFIC

RECOGNIZER PERFORMANCE

Recognition performance can be broken down into two areas: recognition rate and recognition accuracy. Analysis from cockpit recorder data and pilot comments indicate that the VIS demonstrated real-time performance, outputting recognition results on a node by node, or to put it more clearly, a word by word basis, as the command was being spoken.

Recognition accuracy results were mixed. The poor recognition performance of the keyword *VIC-ON* led to a significant amount of pilot frustration. During vocabulary development, a high confidence level was designed into the keyword. That is to say, the required minimum distance score for *VIC-ON* was higher than for all the other words in the vocabulary. What this meant to the pilot operating the system was that the keyword was much less tolerant to variations in speech pronunciation. Interestingly, cockpit recorder data showed that the evaluation pilots tended to say the *VIC-ON* with much more variability than any other word in the vocabulary. The situation was exacerbated further by the evidence that when the keyword was not recognized the first time, the pilots tended to change the way they said the keyword, usually faster and louder, which in all cases was unsuccessful.

Recognition accuracy results are shown in table II. Since the keyword problem presented a special case, recognition accuracy for individual functions was tabulated only for those cases in which the keyword was recognized and the system was activated to listen to the utterance and attempted a pattern match.

Table II
RECOGNITION ACCURACY

Category of Functions	Number of Commands	Number of Mis-Recognitions	Accuracy
All Functions	1084	529	51.2 %
All Functions Activated By Keyword	553	24	95.7%
Communications	171	2	99.1%
Waypoint Selection	52	0	100.0%
EHSI	127	14	90.0%
Stores	36	2	94.4%
Timer	24	1	95 2%
I-P Overfly Update	9	0	100.0%
Target Of Opportunity	12	3	75.0%
Status Report	38	0	100.0%
ECM	21	0	100.0%
Designate	10	1	90.0%
Air-To-Ground	26	1	96.2%
Menu	27	0	100.0%

Note: Individual function accuracies calculated only for cases where keyword was recognized.

EFFECT OF PILOT STRESS ON SPEECH PRODUCTION

A subjective assessment was conducted of the effects of pilot stress on speech production. Analysis of cockpit recorder data indicated high stress levels caused the evaluation pilots to change the way they pronounced words and phrases. High pilot stress was caused by two factors: 1) high cockpit workload and 2) frustration caused by having to repeat voice commands due to VIS recognition errors.

Generally, high stress had the effect of increasing the frequency and volume of speech, which caused distortions in the signal spectrum. This was seen in the cockpit as the pilot raising his voice and increasing his speed of speech. Qualitatively, most recognition errors occurred during high workload phases of flight, particularly during the CAS flights. High stress had the most detrimental effect on the recognition of the keyword *VIC-ON*, which by design was much less tolerant to speech variability than other vocabulary words.

VOCABULARY

The evaluation pilots reported that the vocabulary syntax was cumbersome and difficult to use, as well as confusing, unnatural and inflexible. The vocabulary syntax was designed to mimic the pushbutton switchology logic of the current AV-8B cockpit. This resulted in some long commands which were in some cases difficult for the pilots to remember. For example, to change the target elevation for a waypoint offset point, the pilot was required to say *VIC-ON EHSI DATA OFFSET ELEVATION FIVE ZERO ENTER*. The evaluation pilots commented that it would have been much easier and more natural to access the desired function directly. For the above example the pilot might say *TARGET ELEVATION FIVE ZERO ENTER*. Both evaluation pilots strongly indicated a desire for shorter, more direct commands.

In some cases the evaluation pilots called the same function or control by two different names. Cockpit recorder data indicated that in some instances the pilot's referred to a preset radio frequency as *button* at other times *channel*. The proper vocabulary word was *button*. When the pilot commanded a radio frequency change and used the word *channel* in place of *button*, the command was of course, mis-recognized. During high workload tasks this led to some confusion and frustration on the part of the pilots. Other similar cases involved the pilots saying *mark* instead of *hack* for timer functions, *TOO* instead of *Target Of Opportunity*, and *82 high drag* or *low drag* vice *82 high* or *low*. Both pilots indicated in their postflight debriefs a strong desire for multiple words for functions that are commonly referred to by a different various names.

COMMUNICATIONS FUNCTIONS

Communications functions received favorable comments from the evaluation pilots. Recognition accuracy of communications functions was close to 100%. Both pilots indicated that use of communications functions significantly decreased pilot workload, both airborne and while on deck. During CAS mission preparation, when a large number of frequencies had to be programmed into the radio before takeoff, the pilots commented that voice entry of the data made data entry faster and easier.

In addition to the problem mentioned earlier concerning pilots substituting *channel* for *button* when commanding frequency changes, both pilots frequently violated vocabulary syntax by saying, for example, *button twenty-six* instead of *button two six*, which resulted in a mis-recognition of the command. Both pilots indicated a strong desire to be able to call each button number in either fashion. Pilot comments also indicated a strong desire for the capability to command tactical radio frequencies changes by the color code assigned to those frequencies in the mission planning document, since that was how the TAC(A) and FAC(A) referred to tactical frequencies.

WORKLOAD REDUCTION

A subjective evaluation of workload reduction was conducted during all phases of flight. Pilot comments indicated that the VIS significantly reduced pilot workload. Areas which demonstrated the greatest decrease in pilot workload included: 1) entry of communication, navigation and weapons programming data on the ground and in flight, 2) selecting waypoints in flight, and 3) selecting display pages such as the EHSI, Stores, ECM and Data. Both pilots strongly endorsed the status report functions, commenting that these functions greatly decreased cockpit workload and increased pilot flexibility. Both pilots suggested that VIS technology demonstrated excellent potential to further reduce pilot workload and provide the pilot even greater flexibility than was demonstrated during this evaluation. Although no flights were flown at night, both pilots agreed that the communications, navigation, weapons programming, waypoint selection, and display page functions, as currently implemented, would significantly reduce pilot workload during night low level attack missions, and that the greatest degree of pilot workload reduction would be experienced during all phases of night flight, when visual cues are significantly reduced.

VI. DISCUSSION

RECOGNIZER PERFORMANCE

The requirement to activate the VIS by voice posed some unique challenges to the design of the system. In effect, the VIS was required to listen constantly until it recognized the keyword *VIC-ON*, at which time it would act on the phrase that followed. In order to properly mechanize the keyword function, noise rejection templates were produced which contained small word parts (phonemes). The VIS, which was constantly listening to everything the pilot said, would match the audio input to both the keyword template and the noise rejection templates. The pattern match with the highest minimum distance score was accepted by the system as the uttered phrase, which in this case was either the keyword *VIC-ON* or random speech and noise.

The poor recognition accuracy of the keyword can be attributed in part to greater speaker variability coupled with a higher required minimum distance score for acceptance of the utterance as the keyword. However, the major contributing factor appears to be the noise rejection templates. It is important to reiterate that the VIS compares the score of the match of utterance with the keyword template against the score the utterance with the noise rejection templates. The highest scoring match is accepted by the system. For example, if the pilot coughs or says *rodger* the system will recognize that utterance as noise, not the keyword, and reject it.. The noise rejection templates were made up of the smallest possible word parts, which enabled the system to match and reject random speech and noise very well. However, it appears that when the keyword was spoken with even seemingly normal variations in pronunciation, the VIS was able to produce a better match of the utterance against the noise rejection templates (a concatenation of phonemic word parts) than against the keyword template. The result of this was seen in the cockpit as a mis-recognized, or more accurately, an ignored command.

Two solutions to this problem readily present themselves. A combination of lowering the required minimum distance score and/or adjusting the size of the word parts that make up the noise rejection templates is one approach. However, implementing a manually activated VIS switch and eliminating the requirement for noise rejection templates appears to be a more straight forward and simpler approach.

Reducing the minimum required distance score for keyword recognition would intuitively increase recognition accuracy. However, as the required score is lowered, it is reasonable to expect the probability of false recognition of *VIC-ON* to increase. This would be seen in the cockpit as the system activating on its own. Keep in mind that once the system is activated, it is actively listening to audio input and attempting to make a match. This would result in false commands being executed by the system.

Increasing the size of the noise rejection template word parts would have much the same effect. As the word parts are increased in size, the ability of the VIS to accurately model random speech and noise would correspondingly decrease. It is important to note that the VIS in effect builds a model of the utterance out of a concatenation of noise rejection template word parts. The larger the word parts, the more crude the model, and consequently, the lower the distance score of the match of the utterance against those noise rejection templates. Obviously, the lower the noise rejection template score is, the higher the probability that the keyword template match will be accepted, even though the utterance may not have been the keyword VIC-ON.

The task then, in this approach, is to adjust the size of noise rejection template word parts in combination with setting the proper minimum required distance score for the keyword. Undoubtedly, given enough research and experimentation, this approach could significantly improve the recognition accuracy of the keyword. However, it appears that this approach would be both difficult and time-consuming, and would not guarantee 100% system activation.

The simpler and more reliable approach to system activation would be incorporation of a manually activated VIS switch. In this approach, the system would not be listening unless manually activated; hence the noise rejection templates would not be required. The pilot would have a positive tactile cue of system activation through the depression of the switch. A manually activated switch would ensure 100% system activation, vice the 51% activation rate experienced with the voice activated switch during this evaluation. It is worth noting that with 100% system activation rate, the overall recognition accuracy of the VIS would have been 95.7%. An additional advantage of a manually activated switch is that the pilot would not be required to preface each command with a keyword; hence one very unnatural word would be eliminated from all VIS commands.

Location of the VIS switch is very important, and without question the switch should be easily accessible by the pilot at all times. Placement of the switch on the control stick or throttle would seem the most logical choice and would follow the HOTAS concept already employed in the aircraft.

TIME-CRITICAL FUNCTIONS

Time-critical functions included the navigation overfly update functions and the timer functions. The VIS was mechanized so that when the pilot verbally activated the function, the system would continue to listen for the word *execute* or *hack* before executing the function. The time delay built into these functions was 30 seconds. During the delay between the preparatory command and the execute command the system was listening to all audio input and modeling and rejecting random noise through the noise rejection templates.

The noise rejection feature worked very well and no false execution commands were recognized during the flight evaluations, however, pilots report several frustrating cases of having the system time out before they gave the execute command. Once again, the concept of a manually activated VIS switch lends itself well to the solution of this problem.

The manually activated switch could be mechanized so that the pilot would be able to call up the timer or waypoint overfly function verbally, however, the execute command would be given manually through the action of the pilot releasing the VIS switch when ready to execute the command. This type of implementation would allow the pilot to take as much time as would be needed to maneuver and execute the function.

STATUS REPORT FUNCTIONS

Test results indicated that the VIS enabled access to information that was not available to the pilot through other cockpit displays, as well as provide critical aircraft altitude and fuel status reports while the pilot's eyes were fixed on a target or object outside the cockpit. The weapons system status report was particularly useful and clearly indicated the great potential of voice interactive technology in easing pilot workload.

Status report functions could be easily expanded to include reporting other critical aircraft parameters, such as airspeed and aircraft height above ground, as well as range and bearing to selected ground targets and enemy fighter aircraft. Systems status report functions could also be expanded to include the proposed radar system, the air-to-air weapons system, and the electronic countermeasures system. Even seemingly ordinary checklists, such as the landing checklist, could be voice interactive. For example, one method of employing VIS technology to the landing checklist would be to mechanize the system so that when the landing gear handle was placed in the down position, the VIS would automatically check the position of all required systems and aircraft components for landing. After running through the entire checklist, the system would automatically report *landing checklist complete* or report what subsystems or components were not in the proper status. Other checklists could be mechanized in similar fashion, giving the pilot much greater flexibility and significantly reducing cockpit workload.

TEMPLATE DEVELOPMENT

The test results indicated that the majority of mis-recognition errors occurred when the pilot was under a high level of stress. High pilot stress levels were caused by high pilot workload and pilot frustration with having to repeat commands that were initially ignored or mis-recognized. Conversely, very few errors occurred when the pilots experienced low to moderate stress.

The literature review discussed the fundamental changes to pilot speech production that occur when pilots experience high stress levels. One method of improving recognition of speech uttered under high stress would be to average these speech patterns into the voice templates during the template development process. In order to obtain accurate samples of pilot speech under high stress, it would be important to simulate as closely as possible the actual flight conditions in which the pilot would be expected to utter the command.

The 2F99 Harrier Visual Flight Simulator provides a high fidelity simulation of all phases of AV-8B flight, and would be a logical choice for training voice templates. The 2F99 features an actual AV-8B cockpit mounted on a motion-based platform, surrounded by a 270 degree visual scene. During the simulator flight, pilots wear their personal flight gear, including their helmet and oxygen mask/microphone.

During the proposed template training period the pilot would fly a profile that would encompass all phases of flight that the VIS is designed to operate in, including ground pre-flight preparation.. Samples of each word would be taken for each phase of flight. In this manner, at least one sample of each word uttered at low, moderate and high pilot stress levels would be averaged into the voice templates. Requiring each possible command to be spoken at least once during the simulation flight would ensure that coarticulation effects would be taken into account and their effect minimized.

An additional benefit of developing templates in the flight simulator would be the more natural and more representative speech samples that would be obtained. A certain degree of artificiality could be reasonably expected by placing a pilot in front of a computer in a neutral room and asking him to produce speech samples just as he would in flight. It is logical to assume that the more realistic the training atmosphere is, the more representative the speech samples will be.

At least one 2F99 flight simulator is located at each Harrier operating base, and minimal modifications to the existing simulator would be required to accommodate the VIS training system. The simulator computer software could be easily modified to guide the pilot through a template training tutorial. Additionally, incorporating VIS capability into the operational flight simulator provides new pilots the opportunity to train and become familiar with the VIS.

VOCABULARY

Test results indicated that the vocabulary was cumbersome, inflexible, and unnatural. The vocabulary selected for the evaluation was in fact rigid and highly structured, mainly due to the approach taken to design the vocabulary to emulate the pushbutton logic of the current Harrier controls and displays. As was quickly obvious from the first test flight, pilots do not say commands in quite the same way that they select commands manually. What seemed natural when activating controls manually proved to be very unnatural when activating controls verbally.

Part of the reason for these interesting results may stem from the fact that the test pilots used in this evaluation were both very experienced with the manual switchology of the cockpit; so experienced, in fact, that they very likely had committed to memory almost all the various pushbutton sequences necessary to activate different functions. It is not unreasonable to assume that, when activating a function manually, they executed the required sequence automatically, almost subconsciously. If this line of reasoning is followed, then it would indicate that these pilots were at a disadvantage when attempting to activate controls verbally, and this may have skewed their comments on the usefulness of the vocabulary a bit more to the unfavorable side than was warranted.

It may be argued that, given enough training with the vocabulary, Harrier pilots could become proficient in its use. However, this approach would not eliminate some of the unnatural words and phrases in the vocabulary, most notably VIC-ON. This fact aside, it would be prudent for the system designer of a cockpit interactive system to make the vocabulary as natural and flexible as possible, for the sake of pilot workload reduction. With this in mind, the task becomes one of making the commands short and simple, and the vocabulary natural and easy to remember.

To reiterate, elimination of the voice activated switch would eliminate one very unnatural word (VIC-ON) from all the commands. Commands could be made even smaller by restructuring the vocabulary so that instead of following the manual pushbutton logic of the current cockpit, the pilot could access the desired function with one or at most, two words. The SANDAC computer software could easily be modified to send the proper signal or sequence of signals to the MC to access the desired function. For example, instead of the pilot saying four words (which emulated four keystrokes of the manual pushbutton logic) to access a function, the revised vocabulary would enable him to say just one word. Whereas in the former case each word caused the SANDAC to send one signal to the MC (just as pushing one button caused the DP to send one signal to the MC), now in the latter case, one word would cause the SANDAC to send all four signals, in proper sequence, to the MC. Instead of the pilot being required to say *VIC-ON MENU EHSI DATA* to access the data display, the new vocabulary would allow him to simply say *DATA*..

Flexibility could be built into the vocabulary by designating multiple words for activation of particular functions. For example, radio channels could be accessed by the word *channel* or *button*, both commonly used terms for preset radio frequencies. Even more flexibility could be achieved by enabling a preset control tower frequency to be selected by saying *tower*. Ultimately, a preset frequency such as say, channel twenty nine, could be selected by saying *button twenty nine*, *channel twenty nine*, *button two nine*, *channel two nine*, *button two niner*, or *channel two niner*.

Obviously, radio frequencies are not the only candidates for multiple activation words. Since the VRS-1280 has an operational vocabulary capacity of 800 words, the potential exists to assign multiple commonly used words to various functions without taxing the limits of the recognizer. Naturally, care must be taken to not assign the same word to two different functions. Also very important, vocabulary selection should strive to choose words that are as phonetically dissimilar as possible.

As was documented in earlier studies, vocabulary development is an optimization process. The revised vocabulary described in the preceding paragraphs is much less rigid and structured than the vocabulary used in the evaluation. It is important to remember that the VRS-1280 functions on a finite state grammar network The more structured the vocabulary, the better the real-time operation of the system. The task then, to achieve optimum overall VIS performance, is to strike the proper balance between vocabulary structure, flexibility and size. In doing this, it appears that the designers must assess the pilot population that will use the technology, the amount of training that will be required to achieve proficiency with the new vocabulary, and the effects of the vocabulary size and composition on recognition rate and accuracy.

VII. CONCLUSION

The AV-8B Voice Interactive System demonstrated excellent potential to reduce pilot workload during all phases of flight. Voice-interactive avionics enabled the pilot to control systems and receive feedback without the requirement to visually or manually locate a control or display. Of further benefit was the capability of the system to check and report on avionics and weapons systems status while leaving the pilot's hands free to perform other tasks. Taking into account the problems encountered with activating the system through the voice-activated switch, the VIS demonstrated a level of performance that warrants further development of the system for integration into operational Harrier aircraft.

Integration of a manually activated VIS switch into the system would greatly improve system performance and pilot confidence in the system. A more natural and flexible vocabulary would result in less mis-spoken and mis-recognized commands. Pilot frustration and stress levels would also be correspondingly reduced. Recognition accuracy could also be improved through the use of more realistic vocabulary training, such as use of a flight simulator for vocabulary development and pilot training. Given the proper balance between pilot training requirements and a flexible and natural vocabulary, it would not be unreasonable to expect recognition accuracies of 98-99% when used by the average fleet Harrier pilot.

Quite obviously, a cockpit voice interactive system capable of speaker-independent, unlimited vocabulary operation and natural language understanding would be highly desirable. No doubt but that someday that kind of capability will be demonstrated, how-ever the technology appears to be still quite some years away from the cockpit. On the other hand, speaker-dependent pattern-matching VRS technology has demonstrated a level of maturity that enables limited cockpit voice-interactive capability to be realized today. The Marine Corps has identified a need for systems to reduce pilot workload during tactical flight. Current VIS technology meets that requirement. A cockpit voice interactive control system, including a speaker-dependent pattern-matching VRS, should be further developed and introduced into operational Harrier aircraft as soon as possible.

BIBLIOGRAPHY

BIBLIOGRAPHY

Anderson, T.R., T.J. and R. McKinley (1985). "Issues in the Development and Use of a Speech Recognition Data Base for Military Cockpit Environments." In: Proceedings of Speech Tech'85, New York: Media Dimensions, pp. 172-175.

Aretz, A.J. (1983). "A Comparison of Manual and Vocal Response Modes for the Control of Aircraft Systems." Proc. Human Factors Society 27th Annual Meeting, (Norfolk, Va), pp. 97-101.

Baker, J.K. (1975). "Stochastic Modeling for Automatic Speech Understanding," In. Speech Recognition, Ed. D.R. Reddy. New York: Academic Press. pp. 521-542.

Beek, Bruno, Edward P. Neuberg, and David C. Hodge (1978). "An Assessment of the Technology of Automatic Speech Recognition for Military Applications." In Automatic Speech and Speaker Recognition. Ed. N. Rex Dixon and Thomas B. Martin. New York: IEEE Press.

Bell, D. and R.W. Becker (1983). "Designing Experiments to Evaluate Speech Input/Output Devices and Applications." Speech Technology, 1(4), pp. 70-75, 78, 79.

Bennett, Raymond W., Steven L. Greenspan, Ann K. Syrdal, Judith E. Tschirigi, and John J. Wisowaty (1989). "Speaking To, From, and Through Computers: Speech Technologies and User-Interface Design." AT&T Technical Journal, 70, pp. 17-29.

Bisiani, Roberto (1985). "Computer Systems for High-Performance Speech Recognition." In New Systems and Architectures for Automatic Speech Recognition and Synthesis. Ed. Renato DeMori and Ching Y. Suen. Berlin: Springer-Verlag.

Bridle, J.S. (1978). "An Efficient Elastic-Template Method for Detecting Given Words in Running Speech." In Automatic Speech and Speaker Recognition. Ed. N. Rex Dixon and Thomas B. Martin. New York: IEEE Press.

------------ and M.D. Brown (1982). "A Data-Adaptive Frame Rate Technique and its Use in Automatic Speech Recognition." Proceedings of the Institute of Acoustics Autumn Conference, pp. 167-172.

Bond, Z.S., T.J. Moore and B. Gable (1986). "Some Phonetic Characteristics of Speech Produced In Noise." Journal of the Acoustic Society of America, S49.

----------------------------- and T.R. Anderson (1987). "The Effects of High Sustained Acceleration on the Acoustic-Phonetic Structure of Speech: A Preliminary Investigation." Journal of the American Voice I/O Society, 4, pp. 1-19.

Bush, M.A. and K. Hata (1983). "An Acoustic Analysis of Coarticulatory Phenomena in Connected Digit Strings." Journal of the Acoustic Society of America, 74 (Suppl 1), pp. S14-S15.

Carpenter, B.E. and S.H. Lavington (1973). "Influence of Human Factors on the Performance of a Real-Time Speech Recognition System." Journal of the Acoustic Society of America, 53, pp. 43-45.

Casali, Sherry Perdue, Beverly H. Williges and Robert D. Dryden (1990). "Effects of Recognition Accuracy and Vocabulary Size of a Speech Recognition System on Task Performance and User Acceptance." Human Factors, 32, pp. 183-196.

Chan, C-F., E. Cheung, P. Ramesh, A-C Hsueh, and B. Hanson (1985). "Experimental and Evaluation Procedures for Developing a Robust Noise and Gain Insensitive ASR Algorithm." In: Proceedings of Speech Tech'85, New York: Media Dimensions, pp. 177-179.

Coler, C.R (1982). "Helicopter Speech Command Systems: Recent Noise Tests Are Encouraging." Speech Technology, 1(3), pp. 76-81.

------------- (1984). "In-Flight Testing of Automatic Speech Recognition Systems." In: Proceedings of Speech Tech'84, New York: Media Dimensions, pp. 95-98.

-------------, R.P. Plummer, E.M. Huff and M.H. Hitchcock (1977). "Automatic Speech Recogniton Research at NASA-Ames Research Center." Proceedings of Workshop on Voice Technology for Interactive Real-Time Command/control Systems Applications, Palo Alto, CA: NASA-Ames Research Center, pp. 143-167.

Cooper, M. (1987). "Human Factors Aspects of Voice Input/Output." Speech Technology, 3(4), pp. 82-86.

Cotton, J.C., M.E. McCauley, R.A. North, and M.I. Streib (1983). Development of Speech Input/Output Interfaces for Tactical Aircraft. Dayton, OH: AFWAL TR-83-3073, (AD-A136 485).

Crittenden, Lorri J. Newton C. Ellis and Rodger J. Koppa (1989). "Considerations of Noise for the Use of Compressed Speech in a Cockpit Environment." Proceedings of the Human Factors Society: 33rd Annual Meeting, 1, Santa Monica, CA: Human Factors Society, pp. 38-42.

Dennison, T. (1987). "The Effect of Simulated Helicopter Vibration on the Accuracy of a Voice Recogniton System." Proceedings of the American Helicopter Society Annual Forum and Technical Display, 1, pp. 133-136.

DeSipio, R. and E. Fry (1983). "Avionics System Plays 'Ask and Tell' With Its Operator." Speech Technology, 1(4), pp. 89-93.

Doddington, G. (1986). "Whither Speech Recognition?" In: W.A. Lea (Ed); Trends in Speech Recognition, Apple Valley, MN: Speech Science Publications, pp. 556-561.

Eason, T. (1987). "AV-8B Night Attack Capabilities and Limitations". Naval Air Test Center Report, SA-02R-87, 24 March 1987.

Edgerton, M.J. (1986). "Experiences with Microphones for Voice Input to Automatic Speech Recognizers." Proceedings 1986 Voice I/O Systems Applications Conference (AVIOS), Alexandria, VA, pp.425-430.

Endres, G. (1991). "Towards The 'Intelligent' Aircraft." Interavia Aerospace Review, 46, pp. 57-61.

Flanagan, J.L. (1972). Speech Analysis, Synthesis and Perception. New York: Springer-Verlag.

Fu, K.S. (1974). Syntactic Methods in Pattern Recognition. New York: Academic Press.

Gallant, John (1990). "Voice Recognition System Learns User Vocabularies and Speaking Styles." EDN, 24 May 1990, p.105.

Gordon, Dean F. (1990). "Voice Recognition and Systems Activation for Aircrew and Weapons System Interaction." In Proceedings of the IEEE National Aerospace and Electronics Conference; 21-25 May 1990, 2,pp. 744-748.

Haton, Jean-Paul (1985). "Knowledge-Based and Expert Systems in Automatic Speech Recognition." In New Systems and Architectures for Automatic Speech Recognition and Synthesis. Ed. Renato DeMori and Ching Y. Suen. Berlin: Springer-Verlag.

Henderson, Breck W.(1989). "Army Pursues Voice-Controlled Avionics to Improve Helicopter Pilot Performance." Aviation Week and Space Technology, 22 May 1989, pp. 43-46.

Hitchcock, M.H., and C.R. Coler (1978). "Microprocessor Based Automatic Speech Recognition Systems for Flight Applications." Journal of the Acoustic Society of America, 64 (Suppl 1), p. S73.

Holden, James (1991): ITT Defense Communications. Personal Interview. 10 June 1991.

Holmes, J.N. (1982). "A Survey of Methods for Digitally Encoding Speech Signals." The Radio and Electronic Engineer, 52, pp. 267-276.

Howell, P. (1983). "The Extent of Coarticulatory Effects: Implications for Models of Speech Recognition." Speech Communications, 2, pp. 159-163.

Hunt, M.J., M. Lennig and P. Mermelstein (1980). "Experiments in Syllable-Based Recognition of Countinuous Speech". Proceedings IEEE Conference on Acoustics, Speech and Signal Processing'80 ,Denver: pp. 880-883.

Kersteen, Z.A. and D. Damos (1983). Human Factors Issues Associated with the Use of Speech Technology in the Cockpit. NASA-CR-166548.

Kurzweil, Raymond (1989). "Beyond Pattern Recognition." Byte, Dec 1989, pp. 277-288.

Landell, B.P., R.E. Wohlford and L.G. Bahler (1986). "Improved Speech Recognition In Noise." Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing 86 (Tokyo), Paper 14.14, pp. 749-752.

LaPorte, C.Y. (1985). "A Description of the Architecture of a Voice Input/Output Evaluation System and its Application in a Military Jet Aircraft." In: Proceedings of Speech Tech'85, New York: Media Dimensions, pp. 313-316.

Lea, Wayne A. (1990). "Speech Recognition: What is Needed Now?" In Trends In Speech Recognition. Ed. Wayne A. Lea. Englewood Cliffs, N.J.: Prentice-Hall.

Lee, C.H. (1989). "Applications of Dynamic Programming to Speech and Language Processing." AT&T Technical Journal, 68, pp. 114-130.

----------- and L.R. Rabiner (1988). "A Frame-Synchronous Network Search Algorithm for Connected Word Recognition," Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, New York: pp. 410-413.

Levinson, Stephen E. (1990). "Recognition of Continuous Complex Speech by Machines." Journal of the Acoustic Society of America, 87, pp. 422-423.

--------------- (1977). "Improving Word Recognition Accuracy by Means of Syntax." Record of the IEEE Conference on Acoustics, Speech and Signal Processing 77 Hartford, CT: pp. 479-482.

--------------- and David B.Roe (1990). "A Perspective on Speech Recognition." IEEE Communications Magazine, Jan 1990, pp. 28-34.

--------------- and K.L. Shipley (1980). "A Conversational Mode Airline Information and Reservation System Using Speech Input and Output." Bell System Technical Journal, 59, pp. 119-137.

Lu, D.Z., A.M. Smith and D.G. Childers (1980). Noise Cancelling in Speech Processing: A Prelude to Feature Extraction." Proceedings Fifth International Conference on Pattern Recognition (Miami Beach, FL), pp. 20-23.

Makhoul, J. (1985). "Automatic Speech Recogniton in Severe Environments: A Report on the National Research Council Study." In: Proceedings of Speech Tech'85, New York: Media Dimensions, pp. 230-239.

Malkin, F.J. (1986). "The Effects on Computer Recognition of Speech When Speaking Through Protective Masks." In: Proceedings of Speech Tech'86, New York: Media Dimensions, pp.265-268.

--------------- and T.W. Dennison (1986). "The Effect of Helicopter Vibration on the Accuracy of a Voice Recognition System." Proceedings of the IEEE National Aerospace and Electronics Conference 86 (Dayton, OH), 3, pp. 813-817.

Mariani, J. (1989). "Recent Advances in Speech Processing," Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, S1, pp. 257-286.

McKinley, R.L. and T.J. Moore (1987). "Air Force Requirements for Speech Synthesis in Tactical Aircraft." In: Proceedings of Speech Tech'87, New York: Media Dimensions, pp.151-154.

Melocco, J.M. (1984). "Voice Command on Combat Aircraft." Military Technology, Electronics in Defense, pp.44-46.

Montford, S.J. and R.A. North (1980). "Voice Entry for Reducing Pilot Workload." Proceedings of the 24th Annual Meeting of the Human Factors Society (Los Angeles), pp. 185-189.

Moore, C.A., and R.D. Moore (1987). "Pilot-Aircraft Interface: The Voice Channel." In: Proceedings of Speech Tech'87, New York: Media Dimensions, pp.175-179.

Moore, C.A. and J.C. Ruth (1984). "Use of Voice Recognition Integrated With Aircraft Cockpit Displays." Proceedings of Symposium of the Society for Information Displays.

Moore, C.A., R.D. Moore and J.C. Ruth (1984). "Applications of Voice Interactive Systems--Military Flight Test and the Future." Proceeding of Conference of the American Institute of Aeronautics and Astronautics.

Moore, Roger K. (1985). "Systems for Isolated and Connected Word Recognition." In New Systems and Architectures for Automatic Speech Recognition and Synthesis. Ed. Renato DeMori and Ching Y. Suen. Berlin: Springer-Verlag.

Moore, Thomas J. (1989). "Speech Technology in the Cockpit." In Aviation Psychology. Ed. Richard S. Jensen. Brookfield, VT: Gower Publishing.

-------------------- and Z.S. Bond (1987). "Acoustic-Phonetic Changes in Speech Due to Environmental Stressors: Implications for Speech Recognition in the Cockpit." In: R.S. Jensen, Ed. Proceedings of 4th International Symposium on Aviation Psychology. Columbus, OH: OSU Aviation Psychology Laboratory, pp.77-83.

Morgan, Nelson (1984). Talking Chips: IC Speech Synthesis. New York: McGraw-Hill.

North, R.A. and H. Bergeron (1984). "Systems Concept for Speech Technology Application in General Aviation." Proceedings 6th Digital Avionics Systems Conference, pp.184-189.

------------ and W.A. Lea (1982). Application of Advanced Speech Technology in Manned Penetration Bombers. Dayton, OH: AFWAL-TR-82-3004, (AD-A119274).

O'Shaughnessy, D. (1987). Speech Communication: Human and Machine. Addison-Wesley: Reading, MA.

Paul, D.B., R.P. Lippman, Y. Chen and C.J. Weinstein (1986). "Robust HMM-Based Techniques for Recognition of Speech Produced Under Stress and in Noise." In: Proceedings of DARPA Strategic Computing Program Speech Recognition Workshop, Palo Alto, CA.

Peacocke, Richard D. and Daryl H. Graf (1990). "An Introduction to Speech and Speaker Recognition." Computer, 23, pp. 26-33.

Pisoni, D.B., R.H. Bernacki, H.C. Nusbaum, and M. Yuchtman (1985). "Some Acoustic-Phonetic Correlates of Speech Produced in Noise." In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, pp. 1581-1584.

Pondaco, Dominick (1989): U.S. Army Avionics Research & Development Activity. Personal Interview. 22 Nov 1989.

Poritz, A.B. (1988). "HMMs: A Guided Tour." Proceedings of the International Conference on Acoustics, Speech, and Signal Processing. S-1, pp. 429-440.

Rabiner, L.R. (1989). "A Tutorial on HMMs and Selected Applications in Speech Recognition." IEEE, 77 (2), pp. 257-286.

---------------- and R.W. Schafer (1978). Digital Processing of Speech Signals. Englewood Cliffs, NJ: Prentice-Hall.

---------------- and C.E. Schmidt (1980). "Application of Dynamic Time Warping to Connected Digit Recogniton. IEEE Trans. Acoustics, Speech and Signal Processing, ASSP-28, pp. 377-388.

Rajasekaran, P.K. and G.R. Doddington (1986). "Experiments in Connected Word Recognition of Speech Under Stress." Proceedings 1986 Digital Signal Processing Workshop (IEEE Acoustics, Speech and Signal Processing Society) (Chatham, MA), pp. 1.7.1 - 1.7.2.

-- (1985). "Speech Recognition in the F-16 Cockpit Using Principal Spectral Components. Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing 85, Tampa, FL: pp. 882-885.

-- and J.W. Picone (1986). "Recognition of Speech Under Stress And In Noise." Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing 86, Tokyo: Paper 14.10, pp. 733-736.

Rash, Wayne (1989). "A Helping Hand." Byte, Dec 1989, pp. 129-130.

Rollins, A. and J. Wiesen (1983). "Speech Recognition and Noise." Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing 83,Boston: pp. 523-526.

Roscoe, Stanley N.(1989). "The Eye Prefers Real Images." In Aviation Psychology. Ed. Richard S. Jensen. Brookfield, VT: Gower Publishing.

Rosenhover, F.A. (1987). "AFTI/F-16 Voice Interactive Avionics Evaluation." In: Proceedings of SAE Aerospace Technology Conference and Exposition. Paper 871751.

Ruth, J.C. (1985). "Applications of Voice Interactive Systems." In: Proceedings of Speech Tech'85, New York: Media Dimensions, pp. 223-227.

Schafer, R.W. and L.R. Rabiner (1975). "Digital Representations of Speech Signals." Proceedings of the IEEE, 63, pp. 662-667.

Simpson, C.A. (1991). NASA-Ames Research Center. Personal Interview. 12 July 1991.

-----------------, C.R. Coler and E.M. Huff (1982). "Human Factors of Voice Input/Output for Aircraft Cockpit Controls and Displays." Proceedings of Workshop on Standardization for Speech Input/Output Technology (National Bureau of Standards, Gaithersburg, MD).

-----------------, C.M. Moore and J.C. Ruth (1986). "Testing for Robust Speech Recogniton Performance." Proceedings of the IEEE National Aerospace and Electronics Conference 86 (Dayton, OH), 3, pp. 790-817.

Singer, E. (1981). "The Effects of Microphones and Facemasks on LPC Vocoder Performance.: Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing'81, pp. 1066-1069.

Smith, K. and W.J.C. Bowen (1986). "The Development of Speech Recognition for Military Fast Jet Aircraft." Proceedings of Military Speech Technology, Washington, D.C.

Steuver, J.W. and R.L.Mckinley (1987)."Audio Technology for Aircraft." Proceedings of Military Speech Technology, Arlington, VA, pp. 92-96.

Tarnaud, P. (1986) "Applications of Voice Input/Output in Future Cockpit Design." In: Proceedings of Speech Tech'86, New York: Media Dimensions, pp. 63-65.

Taylor, M.R. (1986). "Isolated Word Recogniton Based on Distinctive Feature Extraction and Dynamic Time Warping." Proceedings IEE International Conference on Speech Input/Output; Techniques and Applications (London), pp. 8-14.

Teja, Edward R. and Gonella, Gary. (1983). Voice Technology. Reston, Va. :Prentice-Hall, 1983.

VanBronkhorst, A. and R. Abraczinskas (1982). "Testing of Voice Control Device for Aircraft Avionics Applications." Proceedings of Workshop on Standardization for Speech Input/Output Technology (National Bureau of Standards, Gaithersburg, MD).

Vemula, R. E. Lee, and G. Boukis (1985). "A Novel Voice Activated Switch (VOX) for High Noise Environments." In: Proceedings of Speech Tech'85, New York: Media Dimensions, pp. 133-136.

Vidulich, Michael A. and Michael R. Bortolussi (1988). "Speech Recognition In Advanced Rotorcraft: Using Speech Controls To Reduce Manual Control Overload." In Proceedings of the American Helicopter Society National Specialists' Meeting: Automation Applications For Rotorcraft. 4-6 April 1988. Atlanta: American Helicopter Society.

Voorhees, J.W. and N.M. Bucher (1985). "Voice/Visual Display Systems in Stressful Environments." In: Proceedings of Speech Tech'85, New York: Media Dimensions, pp. 233-236.

------------------ and Z.A. Kersteen (1983). "The Role of Voice Technology in an Integrated Cockpit System." Proceedings 91st Annual Meeting of the American Psychological Association (Anaheim, CA).

Warner, N.W. and S.D. Harris (1984). "Voice Controlled Avionics: Programs, Progress and Prognosis." In: Proceedings of Speech Tech'84, New York: Media Dimensions, pp. 110-123.

Werkowitz, E.(1984). "Speech Recognition in the Tactical Environment: The AFTI/F-16 Voice Command Flight Test." In: Proceedings of Speech Tech'84, New York: Media Dimensions, pp. 103-105.

Westerhoff, C.J. and L. Reed (1985). "Voice Systems in U.S. Army Aircraft." Speech Technology, 2(4), pp. 43-50.

White, George M. (1990). "Natural Language Understanding and Speech Recognition." Communications of the Association for Computing Machinery, 33, pp. 72-82.

-------------------- (1976). "Speech Recognition: A Tutorial Overview." Computer, 9, pp. 40-53.

-------------------- and Marvin R. Sambur (1980). "Speech Recognition Research at ITT Defense Communications Division." In Trends In Speech Recognition. Ed. Wayne A. Lea. Englewood Cliffs, N.J.: Prentice-Hall.

Williams, C.E. and K.N. Stevens (1981). "Vocal Correlates of Emotional States." In: J.K. Darby, Jr. (Ed), Speech Evaluation in Psychiatry, New York: Grune and Stratton, pp. 221-240.

Williamson, D.T. (1987). "Flight Test Results of the AFTI/F-16 Voice Interactive Avionics Program." In: Proceedings 1987 Voice I/O Systems Applications Conference, San Francisco, CA.

------------------ and D.G. Curry (1984). "Speech Recognition Performance Evaluation in Simulated Cockpit Noise." In: Proceedings of Speech Tech'84, New York: Media Dimensions, pp. 99-102.

Wood, D. (1985). "A Survey of Algorithms and Architecture for Connected Speech Recognition." In New Systems and Architectures for Automatic Speech Recognition and Synthesis. Ed. Renato DeMori and Ching Y. Suen. Berlin: Springer-Verlag.

Wyatt, B. (1983). "Interactive Voice Control: Co-Pilots of the Future." Speech Technology, 2(1), pp. 60-64.

Zolton-Ford, E. (1984). "Reducing Variability in Natural Language Interactions With Computers." In: Proceedings of the Human Factors Society: 28th Annual Meeting, 1, Santa Monica, CA: Human Factors Society, pp.99-102.

APPENDIX A

TABLE OF CONTENTS

VOCABULARY FUNCTIONS AND DEFINITIONS

VIS ACTIVATION/DEACTIVATION

- VIC-ON - Activates VIS; will initiate/interrupt (reinitiate) any process; always an open command. The system will respond with "READY" when activated.
- DISENGAGE - deactivates VIS; this is an optional command - VIS will automatically time-out in 10 sec if no acceptable command is received (specific cases command a 30 sec time-out, others cancel it completely). The system will respond with "DISENGAGE" when it is deactivated.

COMMUNICATION

- COMM-1 - Selects radio no. 1 for frequency change, opens VIS vocabulary page 8.
- COMM-2 - Selects radio no. 2 for frequency change, opens VIS vocabulary page 8.
- BUTTON - Opens data entry page to select preset channel.
- MANUAL - Opens data entry page to select manual frequency.

DATA ENTRY (Vocabulary Page 14)

- ZERO through NINE - Self explanatory
- POINT or PLUS - Interchangeable commands for decimal point (PLUS normally used as division between minutes and seconds when entering time).
- CLEAR - Erases scratch pad, remaining on the data entry vocabulary page.
- ENTER - Completes the line of data entry, closing the data entry page and reopening the previous page. (NOTE - This is the only command which directs VIS to return to a previous vocabulary page. All other actions must be directed from the top down.)

EHSI DISPLAY AND OPERATIONS

- E-H-S-I - Selects the EHSI display and opens VIS vocabulary page 2.
- MARK - Stores aircraft present position on currently selected Mark; steps to next Mark.
- SEQUENCE - Commands EHSI display of up to five above and two below currently selected waypoint
- UPDATE - Boxes UPDT on the EHSI display, calls up ODU options, and opens VIS vocabulary page 5.
- OVERFLY - Following UPDATE, prepares for OVFY on the ODU (awaiting EXECUTE), opening vocabulary page 6.
- EXECUTE - Commands actual OVFY update, displaying position error to the pilot on the scratch pad, and opening vocabulary page 7 for ACPT/REJ.
- HUD - Following UPDATE, selects the designate option on the ODU. The selected waypoint is then visually acquired and designated using the TDC (NOTE: WITH THE HELMET TRACKER, THIS WILL ALSO PERMIT OFF-BORESIGHT UPDATES). Position error is then displayed, and VIS opens vocabulary page 7 for ACPT/REJ.
- ACCEPT - Accepts current position update; if available, repeating accepts altitude update.
- REJECT - Rejects current position and/or altitude update.
- DATA - Boxes DATA; selects data display on DDI; and calls up ODU options. This opens vocabulary page 3, closing page 2.

EHSI DATA OPERATIONS (vocabulary pages 3 and 4)

- ZERO through FOURTEEN - Selects specific WYPT for data entry.
- OFFSET - (1) If commanded from the DATA display, selects ODU options for offset data entry of bearing, range, elevation, and UTM (opening vocabulary page 4, closing page 3).
- ELEVATION - Opens data entry page to enter elevation of the selected waypoint's offset.
- U-T-M - Opens data entry page to enter six or seven digit UTM coordinates (as per current AV-8B mechanism) of the selected waypoint's offset.
- BEARING - Opens data entry page to enter bearing to the selected waypoint's offset.
- RANGE - Opens data entry page to enter range to the selected waypoint's offset.

WEAPON SELECTION AND PROGRAMMING

- STORES - Selects the STORES display and opens VIS vocabulary page 9.
- GUN - Selects GUN and displays the computing (CCIP) A/G gun reticle on the HUD. If another weapon has been previously selected, this will command display of the "Hot Guns" (fixed) reticle on the HUD.
- 82 HIGH - Selects MK-82 High Drag for release (assumes weapon store code has been set in the SMS).
- 82 LOW - Selects MK-82 Low Drag for release (assumes store code preset).
- 76 - "Seventy-six", selects MK-76 practice bomb for release (assumes store code preset).
- 48 - "Forty-eight", selects BDU-48 practice bomb for release (assumes store code preset).
- WEAPON - Selects WPN on the UFC, selects ODU options for QTY, MULT, INT, FUZE. Opens vocabulary page 10 , closing page 9. (NOTE - FUZE selections are available on the ODU, but are not available through VIS for the AV-8B demo.)

WEAPON SELECTION AND PROGRAMMING (continued)

- QUANTITY - Opens data entry page to program quantity of weapons to be released.
- MULTIPLE - Opens data entry page to program multiple.
- INTERVAL - Opens data entry page to program interval.

STOPWATCH

- TIMER - Selects the Stopwatch options on the ODU and opens VIS vocabulary page 11.
- TIME-HACK - Selects the HACK function on the ODU.
- TIME-TO-TARGET - Selects the TTT option on the ODU and opens data entry page to enter time-to-target in minutes PLUS seconds.

WAYPOINT SELECTION (Separate from EHSI)

- WAYPOINT - Opens VIS vocabulary page 12.
- ZERO through FOURTEEN - Immediately commands selection and steering of that WYPT (no ENTER required). This does not change any cockpit displays, other than to provide steering to the new WYPT.
- MARK ONE through THREE - Immediately commands selection and steering to that MARK.
- OFFSET - (2) If commanded from VIS vocabulary page 12, selects steering to the OFFSET of the currently selected WYPT. (NOTE: VIC-ON WAYPOINT TWELVE OFFSET immediately commands steering to the OFFSET for WYPT 12.)

WAYPOINT OVERFLY

- I-P-OVERFLY - Prepares for a WOF on the UFC, opening vocabulary page 6, awaiting an EXECUTE command.
- EXECUTE - Executes the WOF (OVFY UPDATE, selects A/G Master Mode, designates the offset, and provides appropriate attack steering). Position and altitude are then displayed for the pilot, and vocabulary page 7 is opened for ACPT/REJT.
- ACCEPT - Accepts position and/or altitude update.
- REJECT - Rejects position and/or altitude update.

MISCELLANEOUS VIS PAGE 1 FUNCTIONS

- AIR-TO-GROUND - Selects the A/G Master Mode.
- E-C-M - Selects ECM display on the DDI.
- TARGET-OF-OPPORTUNITY - Prepares for TOO on the UFC, opening vocabulary page 6, awaiting an EXECUTE command. (NOTE: This is the only time where EXECUTE does **not** open page 7 for ACPT/REJT.)
- FUEL - Commands verbal readout of total fuel remaining.
- COMBAT - Commands system interrogation of readiness for A/G attack, providing a verbal response (see RESPONSE FUNCTIONS).
- ALTITUDE - Commands verbal readout of current altitude.
- MENU - Commands MENU display on the DDI.
- DESIGNATE - Immediately commands designation of the WYPT or WO/S currently selected.

- FOLLOWING "VIC-ON" "REPORT" - "FUEL", "ALTITUDE", OR "COMBAT" COMMANDS THE SYSTEM TO VERBALIZE EITHER FUEL STATE, ALTITUDE, OR READINESS FOR A/G ATTACK.
- "FUEL"(OR "FUEL STATE") COMMANDS A READOUT OF TOTAL FUEL REMAINING IN THOUSANDS OF POUNDS (E.G., "FUEL 5 POINT 2" EQUALS 5200 LBS REMAINING).
- "ALTITUDE" COMMANDS A READOUT OF CURRENT ALTITUDE. THIS WILL BE VERBALIZED AS "BARO 1,2 POINT 3 ". "RADALT 2 POINT 5" WILL BE VERBALIZED IF VALID RADAR ALTIMITER DATA IS AVAILABLE.
- "COMBAT" COMMANDS THE SYSTEM TO LOOK FOR A/G MASTER MODE AND A/G READY FROM THE STORES MANAGEMENT SYSTEM (WEAPON SELECTED, MASTER ARM ON, AND FUZING SELECTED). IF EITHER ITEM IS NOT AVAILABLE, "AIR-TO-GROUND NOT READY" WILL BE READ BACK TO THE PILOT. "COMBAT" WILL BE VERBALIZED IF THE SYSTEM IS READY TO RELEASE A WEAPON.
- "ERROR" WILL BE VERBALIZED IF THE PILOT'S COMMAND CANNOT BE EXECUTED BY THE MISSION COMPUTER. (NOTE: THIS IS EQUIVALENT TO THE SCRATCH PAD FLASHING WITH AN ENTRY ERROR. VIS IGNORES WORDS NOT FOUND ON AN "OPEN" VOCABULARY PAGE. "ERROR" IS DIRECTED BY THE MC IF IT CANNOT EXECUTE A WORD PASSED BY VIS.)
- "READY" IS VERBALIZED WHEN THE SYSTEM IS ACTIVATED WITH "VIC-ON".
- "DISENGAGE" IS VERBALIZED WHEN THE SYSTEM IS DEACTIVATED (EITHER COMMANDED OR AFTER TIME-OUT).

LIST OF VALID VIS PHRASES

COMM

1. VIC-ON COMM-1 BUTTON EIGHT ENTER.
2. VIC-ON COMM-1 MANUAL THREE TWO FOUR POINT NINE ENTER/
3. VIC-ON COMM-2 BUTTON TWO TWO ENTER.
4. VIC-ON COMM-2 MANUAL TWO FOUR FIVE POINT SIX ENTER/
5. VIC-ON COMM-2 MANUAL THREE FIVE NINE POINT ONE (Pause 5) ENTER.
6. VIC-ON COMM-2 MANUAL THREE FIVE NONE POINT ONE (Pause 5) CLEAR THREE FOUR NINE POINT ONE (Pause 2) ENTER.
7. VIC-ON COMM-1 MANUAL THREE TWO (Pause 5) DISENGAGE.
8. VIC-ON COMM-1 BUTTON FOUR (Pause 2) VIC-ON COMM-1 BUTTON SIX ENTER.
9. VIC-ON COMM-2 MANUAL TWO EIGHT EIGHT (Pause 5) POINT SIX ENTER.
10. VIC-ON COMM-1 (Pause 2) VIC-ON COMM-2 BUTTON SEVEN ENTER.

NOTE: (Pause x) indicates an approximately x second silent pause. Breathing sounds should continue, clearing your throat could be added as a variation.

E-H-S-I

1. VIC-ON E-H-S-I UPDATE OVERFLY (Pause 5) EXECUTE.
2. VIC-ON UPDATE ACCEPT (Pause 5) ACCEPT.
3. VIC-ON E-H-S-I (Pause 5) UPDATE (Pause 5) OVERFLY (Pause 5) EXECUTE.
4. VIC-ON UPDATE REJECT (Pause 5) REJECT.
5. VIC-ON E-H-S-I UPDATE OVERFLY (Pause 2) EXECUTE.
6. VIC-ON UPDATE ACCEPT (Pause 2) REJECT.
7. VIC-ON E-H-S-I UPDATE GPS.
8. VIC-ON UPDATE ACCEPT.
9. VIC-ON E-H-S-I UPDATE GPS.
10. VIC-ON UPDATE REJECT.
11. VIC-ON E-H-S-I (Pause 5) MARK.
12. VIC-ON E-H-S-I SEQUENCE DISENGAGE.
13. VIC-ON E-H-S-I SEQUENCE (Pause 5) SEQUENCE.
14. VIC-ON E-H-S-I DATA (Pause 5) THIRTEEN OFFSET (Pause 5) U-T-M THREE ONE TWO EIGHT ONE SEVEN FIVE ENTER.
15. VIC-ON E-H-S-I DATA (Pause 2) OFFSET RANGE SEVEN POINT TWO ENTER (Pause 2) BEARING ZERO FOUR FIVE ENTER (Pause 2) ELEVATION ONE SIX TWO ZERO ENTER (Pause 2) DISENGAGE.
16. VIC-ON E-H-S-I THREE OFFSET (Pause 2) VIC-ON E-H-S-I DATA (Pause 2) THREE OFFSET U-T-M ONE TWO FIVE ONE NINE ZERO (Pause 5) ENTER (Pause 2) ELEVATION SIX ZERO ENTER.
17. VIC-ON E-H-S-I DATA FOURTEEN OFFSET (Pause 2) RANGE ONE THREE POINT FIVE ENTER (Pause 2) BEARING TWO EIGHT ZERO (Pause 2) CLEAR (Pause 2) TWO EIGHT FIVE ENTER (Pause 2) ELEVATION THREE FOUR ZERO ENTER (Pause 2) DISENGAGE.

STORES

1. VIC-ON STORES EIGHTY-TWO LOW WEAPON (Pause 5) QUANTITY SIX ENTER (Pause 2) MULTIPLE TWO ENTER (Pause 2) INTERVAL FOUR ZERO ENTER (Pause 2) DISENGAGE.
2. VIC-ON STORES SEVENTY-SIX.
3. VIC-ON STORES FORTY-EIGHT (Pause 2) WEAPON (Pause 2) QUANTITY ONE (Pause 2) ENTER (Pause 2) MULTIPLE ONE (Pause 2) ENTER.

STORES (cont)

4. VIC-ON STORES WEAPON (Pause 5) DISENGAGE (Pause 2) VIC-ON STORES EIGHTY-TWO HIGH (Pause 2) WEAPON QUANTITY FOUR ENTER MULTIPLE TWO ENTER (Pause 2) INTERVAL ONE ZERO ZERO ENTER (Pause 2) DISENGAGE.
5. VIC-ON STORES EIGHT-TWO HIGH (Pause 2) GUN.
6. VIC-ON STORES WEAPON (Pause 2) QUANTITY SIX ENTER MULTIPLE ONE ENTER INTERVAL ONE TWO ZERO ENTER (Pause 2) INTERVAL ONE ZERO ZERO ENTER (Pause 2) DISENGAGE.
7. VIC-ON STORES WEAPON INTERVAL ONE ZERO ENTER.

TIMER

1. VIC-ON TIMER (Pause 2) HACK.
2. (Read very rapidly) VIC-ON TIMER HACK.
3. VIC-ON TIMER T-L-T ONE SIX TWO ZERO PLUS ZERO ZERO ENTER DISENGAGE.
4. VIC-ON TIMER (Pause 5) HACK. VIC-ON TIMER TIME-TO-TARGET SIX PLUS ZERO ZERO ENTER (Pause 2) C-S-T.
5. VIC-ON TIMER (Pause 2) T-O-T ZERO NINE FOUR FIVE (Pause 5) CLEAR ZERO NINE FIVE FIVE PLUS THREE ZERO ENTER.
6. VIC-ON TIMER C-S-T DISENGAGE.

WAYPOINT

1. VIC-ON WAYPOINT SIX.
2. VIC-ON WAYPOINT ELEVEN OFFSET.
3. VIC-ON WAYPOINT SEVEN (Pause 2) OFFSET.
4. VIC-ON WAYPOINT OFFSET.
5. VIC-ON WAYPOINT MARK-2.

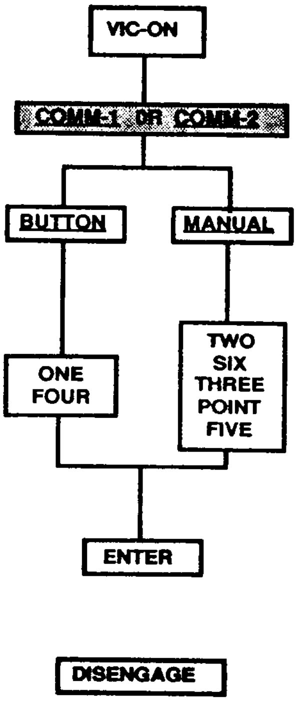
I-P OVERFLY

1. VIC-ON I-P OVERFLY (Pause 5) EXECUTE.
2. VIC-ON UPDATE (Pause 2) ACCEPT (Pause 2) ACCEPT.
3. VIC-ON I-P OVERFLY (Pause 10) EXECUTE.
4. VIC-ON UPDATE ACCEPT REJECT.
5. VIC-ON I-P OVERFLY (Pause 5) EXECUTE.
6. VIC-ON UPDATE REJECT (Pause 5) REJECT.

MISC.

1. VIC-ON MENU DISENGAGE.
2. VIC-ON AIR-TO-GROUND.
3. VIC-ON DESIGNATE (Pause 20 DISENGAGE.
4. VIC-ON E-C-M.
5. VIC-ON TARGET-OF-OPPORTUNITY (Pause 10) EXECUTE.
6. VIC-ON REPORT FUEL.
7. VIC-ON REPORT ALTITUDE.
8. VIC-ON REPORT COMBAT.

VOCABULARY PATH EXAMPLES



ONLY VOCABULARY ON PAGE 1 (OR 15 OR 16 AS REQUIRED BY CURRENT COCKPIT DISPLAYS) PASSED TO MISSION COMPUTER.

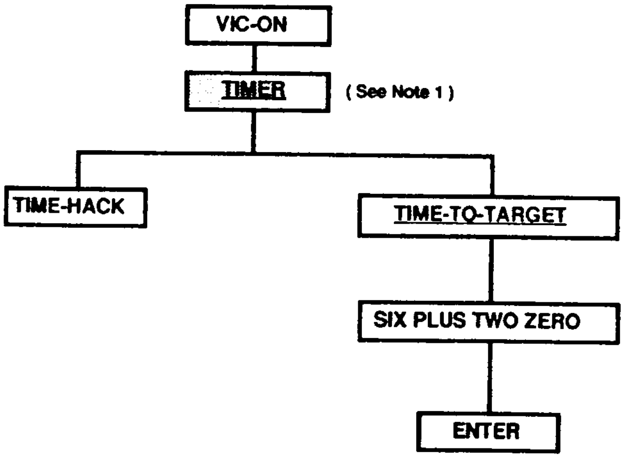
COMM-1 OR *COMM-2* OPENS PAGE 8, CLOSES PREVIOUS PAGE. VIS WILL NOT RESPOND TO ANY COMMANDS OTHER THAN THOSE ON PAGE 8 UNLESS THAT COMMAND IS PRECEEDED BY *VIC-ON* .

BUTTON OR *MANUAL* OPENS PAGE 14, PAGE 8 IS CLOSED.

NUMBERS APPEAR ON SCRATCH PAD AS SPOKEN.

ENTER COMPLETES LINE OF DATA ENTRY AND CLOSES PAGE 14, REOPENING PAGE 8. (*CLEAR* ERASES THE SCRATCH PAD, REMAINING ON PAGE 14.)

DEACTIVATES VIS. SYSTEM WILL AUTOMATICALLY TIME OUT (DE-ACTIVATE) IN 10 SECONDS IF NO ADDITIONAL COMMANDS RECEIVED.

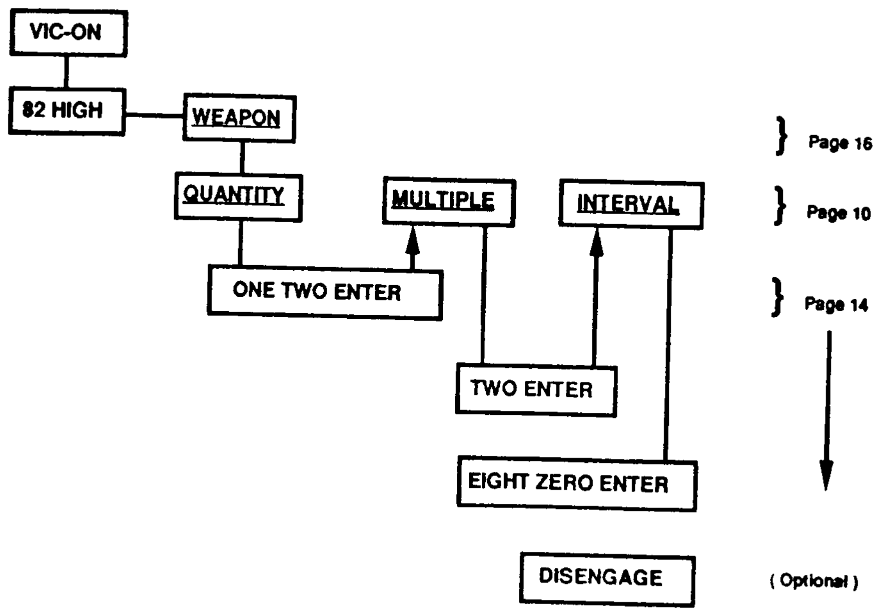


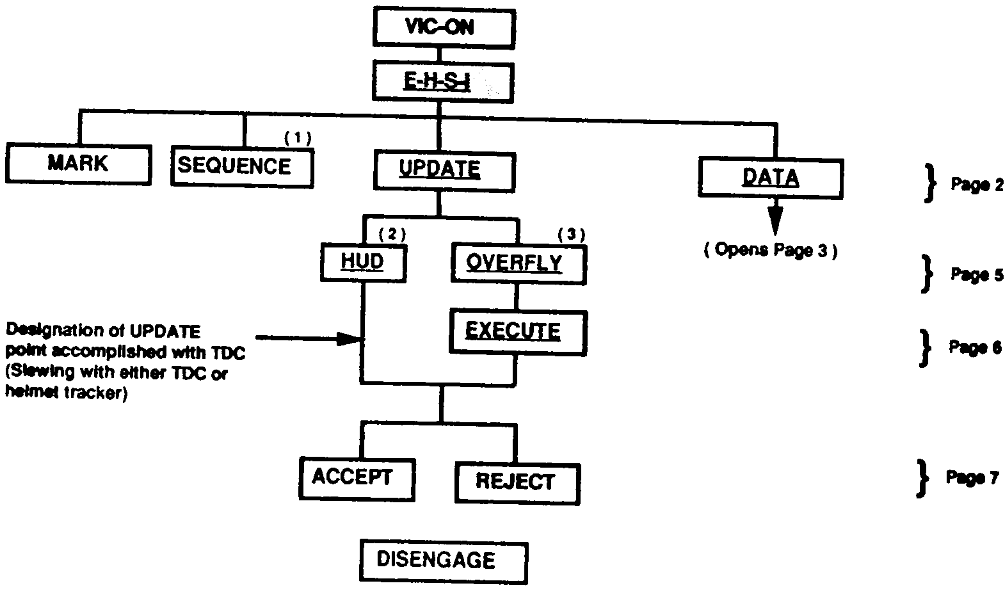
OPENS PAGE 11 AND CLOSES PAGE 1.

OPENS PAGE 14 (Data Entry), CLOSES PAGE 11. NOTE: EXAMPLE SHOWS 6 MINUTES PLUS 20 SECONDS. (*PLUS* IS INTERCHANGEABLE WITH *POINT* .)

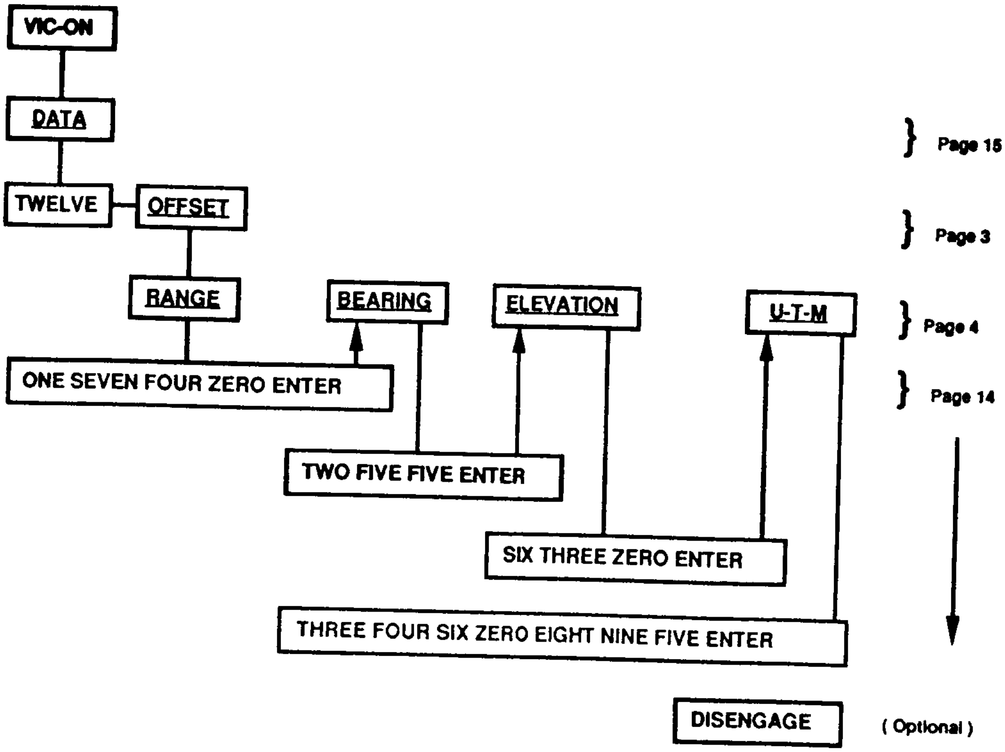
ENTER COMPLETES THE DATA ENTRY, CLOSES PAGE 14, AND REOPENS PAGE 11.

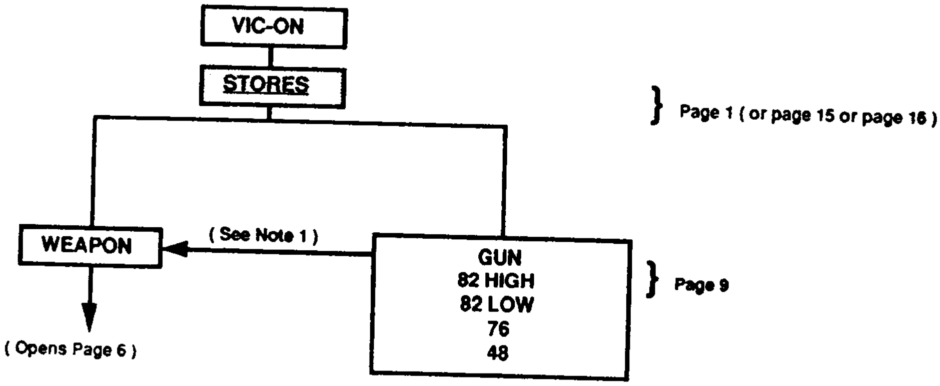
Note 1: Automatic time-out disabled. Pilot must say *DISENGAGE* when process completed.





Notes: (1) Selects or deselects (boxes or unboxes)
(2) Time-out cancelled
(3) Time-out delayed to 30 seconds





Note 1: If the pilot selects WEAPON (WPN) with pushbutton to initiate weapon programming without first selecting a weapon, the ODU will flash. He resolves this by then selecting a specific weapon. Mechanized as shown, he must say *VIC-ON ... 82 LOW ... WEAPON* to resolve.

APPENDIX B

SOFTWARE INTEGRATION TEST

1. EHSI

NOTE: If command not followed by another command within 10 seconds, the VIS will DISENGAGE.

Say VIC-ON	"READY" message given.
Say EHSI	The DDI will show the navigation EHSI display.
1.1	
Say MARK.	Aircraft present position stored in currently selected mark; steps to next mark. VIS automatically disengaged, "DISENGAGE" message given.
1.2	
Say SEQUENCE	If sequence already selected then: - removes box from SEQ legend - removes two prior and next five waypoints from EHSI display, else Boxes "SEQ" legend on EHSI display. Selected waypoint plus two prior and next five waypoints are displayed. Smaller circles and numbers displayed ENDIF. VIS automatically disengaged, "DISENGAGE" message given.
1.3	
Say UPDATE	UPDT box on EHSI display. TCN, DESG, :OVFY, GPS, REJ options on ODU.
1.3.1	
Say OVERFLY.	
Say EXECUTE	:WYPT, FIX, ACPT, REJ options on ODU. Bearing and range error displayed on S/P. VIS automatically disengaged, "DISENGAGE" message given.
Say VIC-ON.	"READY" message given.
Say UPDATE.	

Say ACCEPT. (REJECT)	If ACCEPT, waypoint symbol moves beneath A/C symbol on DDI. (If REJECT symbols do not change). :ALT, ACPT, REJ options on ODU. Baro altitude error displayed on S/P, if available (if not the MC automatically disengages the VIS: "DISENGAGE" message is given, and box from UPDT legend is removed; UFC/ODU blanked).
Say ACCEPT. (REJECT)	UFC/ODU blanked. Box removed from UPDT legend. VIS automatically disengaged, "DISENGAGE" message is given.
1.3.2	
Say GPS	:GPS on ODU Range and bearing error displayed on S/P. VIS automatically disengaged, "DISENGAGE" message given.
Say VIC-ON.	"READY" message given.
Say UPDATE Say ACCEPT	Correction added to INS present position.
Repeat process and say REJECT.	No change in present position.
	UFC/ODU blanked. Box removed from UPDT legend. VIS disengaged by MC. "DISENGAGE" message given.
1.4	
Say DATA.	Selects data display on DDI, boxes "DATA" and "WYPT" and calls up ODU options (WYPT,:POS, ELEV, DECL, UTM).
Say ZERO.	Selects waypoint #0.
(Say ONE).	(Selects waypoint #1).
. . .	. . .
(Say FOURTEEN)	(Selects waypoint #14).
Say OFFSET	:WO/S, BRG, ELEV, RNG,:UTM options on ODU. UTM option is not available if WYPT, DECL and UTM not entered, and colon will be next to BRG.

1.4.1

Say ELEVATION	:ELEV on ODU.
Say a sequence of numbers. (available from ZERO thru NINE).from ZERO (Valid range: 0 - 25000 ft).	Each number said is displayed on S/P. Numerics on UFC disabled.
Say CLEAR	S/P blanked, VIS ready to enter numbers.
Say a sequence of numbers. (available from ZERO thru NINE). (Valid range: 0 - 25000 ft).	Each number said is displayed on S/P. Numerics on UFC disabled.
Say ENTER	Information displayed on DDI. If data out of range "ERROR" message will be given. VIS ready to accept data entry again.

1.4.2

Say UTM	:UTM on ODU.
Say a sequence of numbers. (from ZERO thru NINE). (Range: input 6 or 7 numbers).	Each number said is displayed on S/P. Numerics on UFC disabled.
Say CLEAR	S/P blanked, VIS ready to enter numbers.
Say a sequence of numbers. (available from ZERO thru NINE).	Each number said is displayed on S/P. Numerics on UFC disabled.
Say ENTER	Information displayed on DDI. Range and bearing recalculated. If data out of range "ERROR" message will be given. VIS ready to accept data entry again.

1.4.3

Say BEARING.	:BRG on ODU.
Say a sequence of numbers (ZERO thru NINE). (Range: 0 - 360.0)	Each number said is displayed on the S/P. Numerics on UFC disabled.
Say CLEAR	S/P blanked.

Say a sequence of numbers (ZERO thru NINE). (Range: 0 - 360.0)	Each number said is displayed on the S/P. Numerics on UFC disabled.
Say POINT.	Point displayed on S/P indicating decimals.
Say CLEAR	S/P blanked.
Say a sequence of numbers (ZERO thru NINE). (Range: 0 - 360.0)	Each number said is displayed on the S/P. Numerics on UFC disabled.
Say POINT.	Point displayed on S/P indicating decimals.
Say one number (0-9)	Decimal displayed on S/P.
Say CLEAR	S/P blanked.
Say a sequence of numbers (ZERO thru NINE). (Range: 0 - 360.0)	Each number said is displayed on the S/P. Numerics on UFC disabled.
Say POINT.	Point displayed on S/P indicating decimals.
Say one number (0-9)	Decimal displayed on S/P.
Say ENTER	Information displayed on DDI. If data out of range "ERROR" message will be given.

1.4.4

Say RANGE	:RNG on ODU.
Say a sequence of numbers. (available from ZERO thru NINE).from ZERO (Range: 0 - 100 NM)	Each number said is displayed on S/P. Numerics on UFC disabled.
Say CLEAR	S/P blanked, VIS ready to enter numbers.
Say a sequence of numbers. (available from ZERO thru NINE).	Each number said is displayed on S/P. Numerics on UFC disabled.
Say ENTER	Information displayed on DDI. If data out of range "ERROR" message will be given. VIS ready to accept data entry again.
Say DISENGAGE	"DISENGAGE" message will be given.

2. RADIOS. COMM1 or COMM2

NOTE: If command not followed by another command within 10 seconds, the VIS will DISENGAGE.

Say VIC-ON	"READY" message given.
2.1	
Say COMM-1	Mode, modulation, SQL, encryption, CD#, options are displayed on UFCS. COMM-1 selected and legend in window is cued.
2.1.1	
Say BUTTON	Ready for preselected channels.
Say one or two numbers (ZERO to NINE) (Range: 1-26).	S/P blanked.
Say CLEAR	Ready to accept new channel entry.
Say one or two numbers (ZERO-NINE).	Preset channel selected.
Say ENTER	If channel out of range "ERROR" message will be given. Selected channel displayed over "COMM1" legend on UFC. VIS automatically disengaged, "DISENGAGE" message will be given.
2.1.2	
Say MANUAL	Displays an "M" over COMM-1 legend.
(Range: 30.00 - 399.975).	
Say numbers (ZERO-NINE)	Numbers displayed on S/P. Last number entered is flashing.
Say CLEAR.	S/P blanked.
Say numbers (ZERO-NINE).	Numbers displayed on S/P.
Say POINT	Decimal point displayed on S/P and flashing.
Say CLEAR	S/P blanked.

Say numbers	
Say POINT	
Say numbers (ZERO-NINE)	
Say CLEAR	S/P blanked
Say numbers (ZERO-NINE)	
Say POINT	
Say numbers	
Say ENTER.	If frequency out of range then "ERROR" message will be given and system ready to input new freq, else freq displayed on S/P and VIS automatically disengaged, "DISENGAGE" message will be given.

2.2

Say COMM-2	COMM-2 selected and cued in window.
Say BUTTON	Ready for preselected channels.
Say one or two numbers (ZERO to NINE) (Range: 1-26).	S/P blanked.
Say CLEAR	Ready to accept new channel entry.
Say one or two numbers (ZERO-NINE).	Preset channel selected.
Say ENTER	If channel out of range "ERROR" message will be given. Selected channel displayed over "COMM2" legend on UFC. VIS automatically disengaged, "DISENGAGE" message will be given.

2.2.2

Say MANUAL	Displays an "M" over COMM-2 legend.
(Range: 30.00 - 399.975).	
Say numbers (ZERO-NINE)	Numbers displayed on S/P. Last number entered is flashing.

Say CLEAR.	S/P blanked.
Say numbers (ZERO-NINE).	Numbers displayed on S/P.
Say POINT	Decimal point displayed on S/P and flashing.
Say CLEAR	S/P blanked.
Say numbers	
Say POINT	
Say numbers (ZERO-NINE)	
Say CLEAR	S/P blanked
Say numbers (ZERO-NINE)	
Say POINT	
Say numbers	
Say ENTER.	If frequency out of range then "ERROR" message will be given and system ready to input new freq, else freq displayed on S/P and VIS automatically disengaged, "DISENGAGE" message will be given.

3. STORES. WEAPONS.

NOTE: If command not followed by another command within 10 seconds, the VIS will DISENGAGE.

Say VIC-ON	"READY" message given.
Say STORES	Stores display on DDI.

5.2.3.1

Say GUN	GUN legend boxed on DDI. Displays CCIP reticle on HUD. If another weapon has been previously selected, this will command display of the "Hot Guns" (fixed) reticle on HUD.

3.2

Say 82 HIGH (EIGHTY-TWO HIGH)	82H boxed on DDI. MODE, FUZE, QTY, MULT, TGT ELEV on DDI. If weapon not available "ERROR" message will be given, VIS ready to accept selection of another weapon.

3.2.1

Say WEAPON	:QTY, MULT, FUZE, INTV on ODU
Say QUANTITY: (1-6)	:QTY on ODU.
Say number (ONE-SIX)	Numbers displayed on S/P.
Say CLEAR.	S/P blanked.
Say number (ONE-SIX)	Number displayed on S/P.
Say ENTER	If quantity out of range then "ERROR" message will be given else data is shown at the right of QTY on DDI.
Say MULTIPLE (1-2)	:MULT on ODU
Say ONE or TWO	Number displayed on S/P.
Say CLEAR.	S/P blanked.
Say number different than ONE or TWO (ZERO-NINE).	Number displayed on S/P.

Say ENTER	Entry flashes on S/P. "ERROR" message given.
Say CLEAR.	S/P blanked.
Say ONE or TWO.	Number on S/P.
Say ENTER.	Number displayed at right of MULT on DDI.
NOTE: In previous section end sequence with QTY - 2 and MULT - 1.	
Say INTERVAL (10-200).	:INTV on ODU
Say numbers (ZERO-NINE)	Numbers displayed on S/P.
Say CLEAR	S/P blanked.
Say numbers (ZERO-NINE)	Numbers on S/P.
Say ENTER.	If interval out of range then entry flashes on S/P and "ERROR" message will be given else numbers are displayed at right of INT on DDI.
Say DISENGAGE.	"DISENGAGE" message will be given.
3.3	
Say 82 LOW (EIGHT-TWO LOW)	82L boxed on DDI. MODE, FUZE, QTY, MULT, TGT ELEV on DDI. If weapon not available "ERROR" message will be given, VIS ready to accept selection of another weapon. If weapon not available "ERROR" message will be given, VIS ready to accept selection of another weapon.
3.3.1	
Say WEAPON	:QTY, MULT, FUZE, INTV on ODU
Say QUANTITY: (1-6)	:QTY on ODU.
Say number (ONE-SIX)	Numbers displayed on S/P.
Say CLEAR.	S/P blanked.
Say number (ONE-SIX)	Number displayed on S/P.

Say ENTER	If quantity out of range then "ERROR" message will be given else data is shown at the right of QTY on DDI.
Say MULTIPLE (1-2)	:MULT on ODU
Say ONE or TWO	Number displayed on S/P.
Say CLEAR.	S/P blanked.
Say number different than ONE or TWO (ZERO-NINE).	Number displayed on S/P.
Say ENTER	Entry flashes on S/P. "ERROR" message given.
Say CLEAR.	S/P blanked.
Say ONE or TWO.	Number on S/P.
Say ENTER.	Number displayed at right of MULT on DDI.

NOTE: In previous section end sequence with QTY - 2 and MULT - 1.

Say INTERVAL (10-200).	:INTV on ODU
Say numbers (ZERO-NINE)	Numbers displayed on S/P.
Say CLEAR	S/P blanked.
Say numbers (ZERO-NINE)	Numbers on S/P.
Say ENTER.	If interval out of range then entry flashes on S/P and "ERROR" message will be given else numbers are displayed at right of INT on DDI.
Say DISENGAGE.	"DISENGAGE" message will be given.

3.4

Say 76 (SEVENTY SIX)	76 boxed on DDI. MODE, FUZE, QTY, MULT, TGT ELEV on DDI. If weapon not available "ERROR" message will be given, VIS ready to accept selection of another weapon.

	If weapon not available "ERROR" message will be given, VIS ready to accept selection of another weapon.

3.4.1

Say WEAPON	:QTY, MULT, FUZE, INTV on ODU
Say QUANTITY: (1-6)	:QTY on ODU.
Say number (ONE-SIX)	Numbers displayed on S/P.
Say CLEAR.	S/P blanked.
Say number (ONE-SIX)	Number displayed on S/P.
Say ENTER	If quantity out of range then "ERROR" message will be given else data is shown at the right of QTY on DDI.
Say MULTIPLE (1-2)	:MULT on ODU
Say ONE or TWO	Number displayed on S/P.
Say CLEAR.	S/P blanked.
Say number different than ONE or TWO (ZERO-NINE).	Number displayed on S/P.
Say ENTER	Entry flashes on S/P. "ERROR" message given.
Say CLEAR.	S/P blanked.
Say ONE or TWO.	Number on S/P.
Say ENTER.	Number displayed at right of MULT on DDI.

NOTE: In previous section end sequence with QTY - 2 and MULT - 1.

Say INTERVAL (10-200).	:INTV on ODU
Say numbers (ZERO-NINE)	Numbers displayed on S/P.
Say CLEAR	S/P blanked.
Say numbers (ZERO-NINE)	Numbers on S/P.

Say ENTER.	If interval out of range then entry flashes on S/P and "ERROR" message will be given else numbers are displayed at at right of INT on DDI.
Say DISENGAGE.	"DISENGAGE" message will be given.

3.5

Say 48 (FORTY EIGHT)	48 boxed on DDI. MODE, FUZE, QTY, MULT, TGT ELEV on DDI. If weapon not available "ERROR" message will be given, VIS ready to accept selection of another weapon.

3.5.1

Say WEAPON	:QTY, MULT, FULE, INTV on ODU
Say QUANTITY: (1-6)	:QTY on ODU.
Say numbers (ONE-SIX)	Numbers displayed on S/P.
Say CLEAR.	S/P blanked.
Say number (ONE-SIX)	Number displayed on S/P.
Say ENTER	If quantity out of range then "ERROR" message will be given else data is shown at the right of QTY on DDI.
Say MULTIPLE (1-2)	:MULT on ODU
Say ONE or TWO	Number displayed on S/P.
Say CLEAR.	S/P blanked.
Say number different than ONE or TWO (ZERO-NINE).	Number displayed on S/P.
Say ENTER	Entry flashes on S/P. "ERROR" message given.
Say CLEAR.	S/P blanked.
Say ONE or TWO.	Number on S/P.

Say ENTER.	Number displayed at right of MULT on DDI.
NOTE: In previous section end sequence with QTY = 2 and MULT = 1.	
Say INTERVAL (10-200).	:INTV on ODU
Say numbers (ZERO-NINE)	Numbers displayed on S/P.
Say CLEAR	S/P blanked.
Say numbers (ZERO-NINE)	Numbers on S/P.
Say ENTER.	If interval out of range then entry flashes on S/P and "ERROR" message will be given else numbers are displayed at right of INT on DDI.
Say DISENGAGE.	"DISENGAGE" message will be given.

4. TIMER

NOTE: If command not followed by another command within 30 seconds, the VIS will DISENGAGE.

Say VIC-ON	"READY" message will be given.
Say TIMER	STPW, RSET, HACK and :TTT (or REAL and :TOT; last selected will be on ODU). If TTT cued S/P will show last entry X.XX-Y.YY with Y.YY counting up if no hack and counting fewer if hack. If TOT cued, S/P shows 00.00.00 or last entry.

4.1

Say TIME-To-TARGET (Range: 0.01 - 59.9) (Minutes . seconds)	TTT cued on ODU. 0.00 - 0.00 or X.XX - Y.YY on S/P.
Say one or two numbers (ZERO-NINE).	Numbers representing minutes displayed on S/P.
Say CLEAR	0.00 - 0.00 or X.XX - Y.YY on S/P.
Say one or two numbers.	Numbers on S/P.
Say PLUS	Point on S/P, separating minutes and seconds.
Say CLEAR	Last TTT on S/P.
Say one or two numbers.	Numbers entered on S/P.
Say PLUS.	Point on S/P.
Say one or two numbers.	Minutes, period and seconds on S/P.
Say CLEAR	Old TTT on S/P.
Say one or two numbers	Numbers on S/P.
Say PLUS	Point on S/P.
Say one or two numbers	

Say ENTER	TOT and CST are calculated. XX.XX - YY.YY or S/P. If no hack said before then XX.XX is number entered and YY.YY is counting upward from 0.00, else YY.YY is counting down and represents X.XX minus time since HACK. Caret and lubber on HUD. If TTT entered out of range entry flashes on S/P and "ERROR" message will be given.
(Say DISENGAGE)	
4.2	
Say T-O-T (00.00.00 - 23.59.59)	TOT cued on ODU, REAL on ODU. 00.00.00 or X.X.XX.XX on ODU depending if TTT or TOT entry had been made.
Say numbers (maximum # of four digits) for hours and minutes.	Numbers displayed on S/P.
Say CLEAR	Old TOT on S/P.
Say numbers (maximum of four digits).	
Say PLUS	Period on S/P separating minutes and seconds.
Say CLEAR	
Say number (four max).	Numbers on S/P.
Say PLUS.	
Say two numbers	Seconds on S/P.
Say CLEAR	Old TOT on S/P.
Say numbers (four max).	
Say PLUS	
Say two numbers	
Say ENTER	If data out of range entry will flash and "ERROR" will be given, else: o TTT and CST are calculated o S/P displays date entered.

4.4

Say HACK	If before saying hack 0.00-0.00 on S/P then after hack 0.00-X.XX displayed where X.XX is time since hack else if Y.YY-Y.YY before hack then Y.YY-X.XX after hack where X.XX is Y.YY minus time after hack.

4.5

Say C-S-T	Z.ZZ-X.XX on S/p where X.XX is time since hack, counting down, and Z.ZZ is airspeed to reach steering point when X.XX is 0.00. If an asterisk is shown on S/P, X.XX is hours and minutes, with the period flashing of 1HZ.

5. WAYPOINT SELECTION

NOTE: If command not followed by another command within 10 seconds, the VIS will DISENGAGE.

Say VIC-ON	"READY" message given.
Say WAYPOINT	Boxes WYPT on EHSI display if present. Bearing bug displayed on HUD. Selected waypoint and ground range in nautical miles is displayed on HUD.
Say ZERO	Waypoint #0 selected. HUD WYPT number is 0.
(ONE)	
Say FOURTEEN	Waypoint #14 selected. HUD WYPT number is 14.
Say OFFSET	If offset entered for selected waypoint, waypoint offset steering is selected. WO/S boxed on EHSI, HUD updates to WO/S steering and range. If no offset, ERROR message given then VIS disengages and "DISENGAGE" message given.

6. WOF (WAYPOINT OVERFLY)

NOTE: If command not followed by another command within 10 seconds, the VIS will DISENGAGE.

Say VIC-ON	"READY" message will be given.
Say IP-OVERFLY	Current WYPT and WO/S symbols on EHSI.
Say EXECUTE	UPDT legend boxed on EHSI/DMT. :ALT, ACPT, REJ options on ODU. Baro altitude error displayed on S/P. (If radar altitude invalid ALT option blanked after one sec and OVFY update option displayed). A/G master mode selected. Data block numbers at entered O/S RNG and BRG. WYPT moves to A/C symbol, O/S moves same amount. ATK line initiates from WYPT to WO/S. WO/S boxed, and O/S dashed circle turns to a dashed diamond. Diamond on HUD. Steering to O/S on HUD. VIS automatically disengaged, "DISENGAGE" message given.
Say VIC-ON	"READY" message given.
Say UPDATE	
Say ACCEPT (REJECT).	:OVFY, ACPT and REJ on ODU. Bearing and range errors on S/P.
Say ACCEPT (REJECT).	Symbols do not change. VIS automatically disengaged, "DISENGAGE" message given.

7. MISCELLANEOUS VIS PAGE

NOTE: If command not followed by another command within 10 seconds, the VIS will DISENGAGE.

Say VIC-ON	"READY" message will be given.
7.1	
Say MENU	Calls up menu display on DDI. VIS automatically disengaged, "DISENGAGE" message given.
7.2	
Say AIR-TO-GROUND	Selects the A/G master mode. VIS automatically disengaged, "DISENGAGE" message is given.
7.3	
Say E-C-M	If ECM display already selected on DDI, brings back the previous display selected, else calls up the ECM display on the DDI. VIS automatically disengaged, "DISENGAGE" message given.
7.4	
Say TARGET-OF-OPPORTUNITY	
Say EXECUTE	Designates point below A/C. A/C position stored on MKX. Calculated ELEV stored (0 if radar altitude invalid). A/G master mode selected. MKX increments on EHSI. TGT DSG on HUD and DDI. ATK line on EHSI except from A/A master mode. If a weapon was not previously selected the TD diamond and attack steering displayed. Attack symbology will not be presented and is cued by four slant lines across the center of the HUD. VIS automatically disengaged, "DISENGAGE" message given.

7.5

Say DESIGNATE

If WYPT or WO/S already designated as a target, it will undesignate it, and diamond or dashed diamond turns to a circle or a dashed circle on EHSI display if selected on DDI. Diamond disappears from HUD. Else it will designate it circle or dashed circle turns to a diamond or dashed diamond on EHSI display. Diamond appears on HUD.

VIS automatically disengaged, "DISENGAGE" message will be given.

7.6

Say REPORT

7.6.1

Say FUEL

"FUEL" message will be given. The pilot will receive "number" "point" "number" indicating thousands of pounds of fuel left. If fuel quantity not available "ERROR" message will be given. VIS automatically disengaged, "DISENGAGE" message given.

Set FUEL data invalid

Say VIC-ON

Say REPORT

Say FUEL

"ERROR" message will be given. VIS automatically disengaged, "DISENGAGE" message given.

7.6.2

Say ALTITUDE

"ALTITUDE" message will be given. Baro altitude will be read out to the pilot in the format: "NUMBER (NUMBER) THOUSAND NUMBER HUNDRED" indicating alt. in feet. If baro altitude not available or invalid, "ERROR" message will be given. VIS automatically disengaged, "DISENGAGE" message will be given.

Set altitude invalid

Say VIC-ON	
Say REPORT	
Say ALTITUDE	"ERROR" message will be given. VIS automatically disengaged, "DISENGAGE" message will be given.

7.6.3

Say COMBAT	System will look for A/G Master Mode, Master ARM and A/G ready from the Mission Computer (weapon selected, weapon programming complete, and fuzing selected). If an item is not available it will be read back to the pilot thru his headset ("AIR-TO-GROUND", "MASTER ARM," "WEAPONS"). "COMBAT" message will be given if the above three criteria have been satisfied. VIS automatically disengaged, "DISENGAGE" message will be given.

APPENDIX C

CAS SIMULATION

- PREPARATION: 1) CONTROL POINTS (CP) AND INITIAL POINTS (IP) ENTERED AS WAYPOINTS WITH UTM CORRELATION FOR IPs
  2) COMM PLAN ENTERED AS PRESET CHANNELS
  3) SMP PROGRAMMED FOR (6) MK-82 SE PILOT SELECTABLE ON PARENT RACKS AND THE GUN

- ALERT: 1) MONITOR DASC ON 318.4
  2) WEAPON PROGRAMS 82H Q6 M2 INT 100 N/T
  82L Q6 M2 INT 50 N

- LAUNCH: 1) DASC DIRECTS TO CP ALPHA WITH INTRUCTIONS TO CONTACT ALPHA 7 BRAVO (TAC(A)) ON GOLD
  2) PILOT CONFIRMS WEAPON STATUS, VIS CALIBRATION, ETC.

- CONTROL POINT: 1) PILOT CONTACTS TAC(A) WHO IMMEDIATELY REQUESTS CONTACT ON PURPLE (334.6, NOT PRESET)
  2) PRELIMINARY INFO EXCHANGED
  3) OVFY UPDATE, REJECT UPDATE
  4) FOLLOWING SHORT LOITER, INSTRUCTIONS TO CONTACT HOTEL 5 SIERRA (FAC(A)) ON GREEN.

- MISSION BRIEF: 1) FAC(A) PROVIDES "9-LINE" BRIEF WITH TTT OF 6+00 DIRECTED TO OFFSET LEFT, TARGET DESCRIPTION REQUIRES INCREASED STICK LENGTH.
  2) PILOT ENTERS TARGET UTM AND ELEVATION AS OFFSET TO ASSIGNED IP (WAYPOINT)
  3) PILOT ENTERS TTT
  4) PILOT MODIFIES WEAPON PROGRAM AS REQUIRED FOR 82H Q6 M1 INT 80
  5) "STANDBY ... 6+00 HACK" (BEFORE AND AFTER ODU STOPWATCH FUNCTIONS TIME OUT)

- ATTACK: 1) PILOT DEPARTS CP AS REQUIRED TO MAKE TTT (USING CS/T)
  2) 200 FT AGL; WOF OVER IP; AUTO RELEASE PLANNED
  3) PILOT CONFIRMS WEAPONS READY, MASTER ARM, ETC WITH "VIC-ON REPORT COMBAT". SYSTEM TO RESPOND WITH "COMBAT" OR -- .
  4) 30 SEC OUT, PILOT MANEUVERS TO OFFSET LEFT, LOOKING RIGHT AS DIRECTED BY HTS CUES TO ACQUIRE TARGET. FAC(A) MARKS TARGET.
  5) PILOT POPS AS REQUIRED, "POPPING"
  6) SAM THREAT REQUIRES CHAFF/FLARES DURING POP
  7) PILOT LOCKS ARBS TO THE TARGET; OR NEAR TARGET AND CONVERTS TO CCIP; "WINGS LEVEL" "CLEARED HOT"
  8) USES HOT GUN
  9) WEAPON RELEASE
- EGRESS: 1) PILOT EGRESSES THROUGH ASSIGNED "RP"
  2) THREAT AVOIDANCE

COMM PLAN

BUTTON	FREQ	COLOR	FUNC	CALL SIGN
17	321.9	GOLD	TAC(A)	ALPHA 7 BRAVO
18				
19				
20				
21				
22				
23				
24		GREEN	FAC(A)	HOTEL 5 SIERRA
25				
26				
	334.6	PURPLE	TAC(A)	ALPHA 7 BRAVO

VITA

Captain Dennis Patrick O'Donoghue USMC was born in Pittsburgh, Pennsylvania on February 28, 1958. He graduated from Seneca Valley High School in June, 1976. The following month he entered the United States Naval Academy and in May, 1980 received a Bachelor of Science degree in Mechanical Engineering. Captain O'Donoghue was commissioned as a Second Lieutenant in the United States Marine Corps on May 28, 1980 and entered U.S. Navy Flight School in March, 1981. He received his wings on 17 September, 1982. After five years of operational flying in the AV-8 Harrier tactical jet aircraft, Captain O'Donoghue was selected to attend U.S. Navy Test Pilot School, Patuxent River, Maryland. He graduated from the fixed wing test pilot course in June, 1989.

Captain O'Donoghue has logged over 2400 flight hours in 33 different type aircraft, including over 1400 flight hours in the AV-8 Harrier. He is currently assigned as a test pilot to the Strike Aircraft Test Directorate, Naval Air Test Center, Patuxent River, Maryland.