NASA CONTRACTER **REPORT** 166548

Human Factors Issues Associated with the Use of Speech Technology in the Cockpit

Z.A. Kersteen
D. Damos

(NASA-CR-166548) HUMAN FACTORS ISSUES ASSOCIATED WITH THE USE OF SPEECH TECHNOLOGY IN THE COCKPIT (Arizona State Univ.) 83 p HC A05/MF A01 CSCL 17B N84-17433
Unclas
G3/32 11655

GRANT NAG2-217
December 1983

NASA

NASA CONTRACTER REPORT 166548

Human Factors Issues Associated with the Use of Speech Technology in the Cockpit

Z.A. Kersteen, M.A.
D. Damos, Ph.D.

Department of Psychology
Arizona State University
Tempe, Arizona 85287

Prepared for
Ames Research Center
Under Grant NAG2-217

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

ABSTRACT

This report reviews the human factors issues associated with the use of voice technology in the cockpit and areas for future research are summarized. The current formulation of the LHX avionics suite is described and the allocation of tasks to voice in the cockpit is discussed. State-of-the-art speech recognition technology is reviewed. Finally, a questionnaire designed to tap pilot opinions concerning the allocation of tasks to voice input and output in the cockpit is presented. This questionnaire was designed to be administered to operational AH-1 pilots. Half of the questionnaire deals specifically with the AH-1 cockpit and the types of tasks pilots would like to have performed by voice in this existing rotorcraft. The remaining portion of the questionnaire deals with an undefined rotorcraft of the future and is aimed at determining what types of tasks these pilots would like to have performed by voice technology if anything was possible, ie. if there were no technological constraints.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge several people who made significant contributions to this report. Clay Foushee, Carol Simpson, George Tucker, and Doug Williams all contributed greatly to the design of the pilot questionnaire. We would like to thank Millard Edgerton for several technical discussions, portions of which were incorporated into this report. Thank you to Chuck Jackson for his creative suggestions concerning the use of speech technology in the cockpit. Finally we would like to thank our NASA technical officer, Dr. James Voorhees at NASA-Ames Research Center, Moffett Field, California.

This work was funded by a grant from NASA-Ames Research Center (NAG2-217).

TABLE OF CONTENTS

INTRODUCTION

Advances in technology, particularly microprocessor technology, continue to broaden the scope of military aircraft missions. Coincident with increased mission complexity and aircraft performance capabilities are increased demands upon the pilot who is required to monitor, manage, and interact with these systems. The computer- driven multifunction display and keyboard is the primary medium of interaction between the pilot and various on board systems in emerging cockpit configurations. The multifunction display can supply vast amounts of information in a relatively small amount of space. However, the multifunction keyboard when it is used alone as a means of interacting with a multifunction display places a heavy burden on the pilot's visual and manual resources. Furthermore, no general guidelines have been developed for information display formats that help the pilot process this information quickly and efficiently. New control/display configurations are needed to fully tap the expanded information retrieval capabilities profferred by emerging microprocessor-based avionics.

The Army's new light helicopter program (LHX) planned for operational use in the mid 1990's will use highly capable digital avionics, which will provide greatly improved performance and mission capabilities relative to existing Army helicopters. In addition the crewsize may be reduced to one. The complexity of this aircraft in terms of mission and system requirements coupled with the one crewmember could be the limiting factors in the successful development of these aircraft.

In deference to the criticality of this issue research is being devoted to the design and optimization of the pilot/aircraft interface in the LHX series of aircraft. Many functions will be automated based on data fusion techniques and the use of artificial intelligence. Moreover, based on the assumption that the pilot's visual input/manual output channels are already overburdened, voice interaction with avionic systems will be implemented. Voice command via automatic speech recognition will provide the means for systems control and interaction without necessitating the use of the pilot's manual control resources. Similarly, the use of speech generation as a means of information display and feedback will reduce the visual processing load.

Speech technology, both recognition and generation, has advanced at an extremely rapid rate in the last decade and is becoming increasingly desirable as a medium of interaction between humans and computers since it is a natural and efficient mode of communication that also frees the hands and eyes for other tasks. The benefits associated with speech technology particularly suggest its use in the helicopter cockpit where visual and manual channel loadings are so high. Optimal use of this technology, however, is dependent upon whether it is allocated to those human tasks that are fatiguing, difficult, and distracting. In essence, the primary consideration governing the integration of speech in the cockpit must be human capabilities and needs. Since speech technology offers a new dimension in human/computer interaction, there is a temptation to use it as a mere replacement for visual/manual operations, such as switching functions. Although speech technology can replace a switch closure, one-to-one replacements of visual and manual operations may not fully exploit the speech interface.

This report will first review the human factors issues associated with the use of voice technology in the cockpit and areas for future research will then be summarized. The current formulation of the LHX avionics suite will be described and the allocation of tasks to voice in the cockpit will be discussed. State-of-the-art speech recognition technology will be reviewed. Finally, a questionnaire designed to tap pilot opinions concerning the allocation of tasks to voice input and output in the cockpit will be presented in the appendix. This questionnaire was designed to be administered to operational AH-1 pilots. Half of the questionnaire deals specifically with the AH-1 cockpit and the types of tasks pilots would like to have performed by voice in this existing rotorcraft. The remaining portion of the questionnaire deals with an undefined rotorcraft of the future and is aimed at determining what types of tasks these pilots would like to have performed by voice technology if anything was possible, ie. if there were no technological constraints.

AUTOMATIC SPEECH RECOGNITION

Although the technology is advancing rapidly, state-of-the-art speech recognition is still in its infancy in many respects. Numerous constraints are placed on the user in terms of the number of words that may be recognized at a time, the speed with which words may be spoken in succession, the permissible variability in the pronunciation of each word, and the amount of preparation time needed to use an automatic speech recognition (ASR) device in an operational environment. However, continuing technological advances suggest that by the time we determine how best to interface ASR and the human, these constraints may no longer be of concern.

Before continuing with a discussion of the more complex issues associated with the use of ASR in the cockpit, a brief functional description of this technology is warranted as is the definition of some of the phraseology.

SPEAKER DEPENDENT VS. INDEPENDENT RECOGNITION

Computer recognition of speech can be classified as either speaker dependent or speaker independent with the former being easier to accomplish than the later. Speaker independent means that the device will recognize words spoken by many different speakers, based on only one set of templates. This type of speech recognition is more difficult to accomplish than speaker dependent recognition since human speech patterns, like fingerprints, are unique to each individual. The trick to accomplishing independent speech recognition is to distill the salient features for each word that are common to every individual's utterance of that word. These "universal" features then comprise the reference template for that particular word. It is readily apparent that reference templates formed and used by only one speaker in a speaker dependent system will be much richer in linguistic content (hence yielding better accuracy) than those templates created for use by many speakers.

Due to state-of-the-art limitations in the creation of independent speech recognition reference templates, these devices are primarily limited to recognition of the digits zero through nine and are further constrained by user dialects. For example, an independent speech recognition device which uses templates formed from typically "southern" speech will not recognize those same words as accurately when spoken with a "northern" accent.

A speaker dependent system requires that each user form one set of templates for each word in the working vocabulary. During the training phase the user repeats each word in the specified vocabulary from one to ten times. The exact number of repetitions is dependent both upon the particular device in use and upon the complexity of the vocabulary. The templates are then maintained in the system memory so that during operational use of the machine each incoming utterance is compared to these reference templates. The template that matches most closely is then chosen as the spoken utterance.

Two distinct approaches to the creation of these reference templates have been adopted. One method averages the repetitions of each word in the vocabulary. Typically, this training method requires three or more repetitions of each word in the vocabulary. The resulting templates then are an "averaged" representation of each word that account for slight variations in the pronunciation of these words. With respect to the number of repetitions needed to create optimal reference templates using the averaging technique, more is not always better. There is a point at which additional repetitions cause the templates to lose their clarity. Generally, the manufacturer will recommend the appropriate number of repetitions. A balance must be achieved between too few repetitions (which yields incomplete templates) and too many repetitions.

The other way in which reference templates are created typically requires only one or two repetitions of each vocabulary word. These templates are maintained separately in memory for comparison.

Poock (1982) has shown that a particular speaker dependent system can achieve a limited degree of speaker independence by having several speakers repeat the vocabulary during one training session. Because the device uses the averaging technique it produces a set of reference templates with speech characteristics representative of each speaker. Thus, several speakers can use the device concurrently without having to load separate templates for each individual.

For the most part, however, optimal performance in terms of recognition accuracy will be obtained when recognition is accomplished by one user at a time, based on his or her own set of reference templates.

DISCRETE VS. CONNECTED/CONTINUOUS WORD RECOGNITION

The next issue of importance with respect to ASR is that of discrete vs. connected or continuous word recognition capabilities. A discrete word recognition device, which is the most common type currently available, will recognize single utterances or short phrases (typically up to 1.5 s without pause) in isolation. The user must pause for a predefined length of time (approximately 200 ms) between each utterance. This pause requirement facilitates the endpoint detection of each utterance.

Connected word recognition allows the user to input a short string of words in a connected fashion. Typically, connected word recognition is used with the digits for entering number sequences such as telephone numbers. Connected word recognition, or high speed voice input capability as it is sometimes called, is just beginning to be available commercially at a reasonable price. Connected word recognition capabilities are still quite constrained with respect to the number and type of words that can be recognized in this manner. Continuous word recognition implies the capability to input an unconstrained number of words in a continuous manner (like conversational speech). No commercially available system yet has this capability, and it will probably not be available in the near future. Both connected and continuous speech recognition are more difficult to achieve than discrete word recognition because of two related problems. First, when speech flows freely in connected form, word boundaries are extremely hard to detect. Second, words distort the pronunciation of adjacent words, a phenomena known as co-articulation. For example, think about saying "Let's go eat". The actual pronunciation is likely to sound similar to "Skweet" (Lea, 1979).

Current connected speech recognition systems deal with the enormous task of sorting through the complexities of conversational speech by limiting the task to the recognition of connected digit strings and to structured command sequences. This structured command language is incorporated into a system by the use of syntax, which represents all the valid word sequences that constitute commands to an ASR system. Syntax structures limit the number of possible words for recognition to those which are valid at that point in the command sequence. For example, syntax structures might be used to aid ASR in the cockpit for a function such as tuning a radio. The recognizer would look for the word "radio" and then look for a string of digits. However, the recognizer would not look for any "nav" functions. The clever use of syntax structures, therefore, limits the number of active word choices at each point in the command sequence. This method is clearly more efficient than choosing among all the words in the vocabulary at all times.

PERFORMANCE MEASUREMENT

There are two types of errors associated with ASR devices. Substitution errors or misses comprise the incorrect recognition of an utterance. For example, the user says "TUNE" and the machine recognizes the word "SLEW." This type of error is by far the most critical in the aircraft environment.

The second type, rejection errors, occur when an incoming utterance fails to match any of the reference templates in memory. Most commercially available ASR devices have a user selectable rejection threshold. This threshold dictates the number of bits that must match between an incoming utterance and a reference template for recognition to occur. A trade-off occurs when selecting a rejection threshold. With a stringent setting, few, if any, substitution errors will occur at the expense of increased utterance rejections. Thus, the user may have to repeat a word several times for classification to occur. With less stringent rejection threshold settings, the machine will attempt to classify all utterances, thereby increasing substitution errors with a concurrent decrease in rejections. An optimal rejection threshold is one in which substitution errors are virtually eliminated while rejections are kept to a minimum. Although substitution errors are clearly the less desirable of the two types of errors, the need to repeat an utterance frequently can be extremely annoying.

A standardized performance metric for the various ASR devices has yet to be accepted. There is currently no generally accepted way to weight the relative seriousness of a substitution error as opposed to a rejection error. Furthermore, a standard method for comparison of ASR devices has yet to be adopted.

SPEECH GENERATION

DIGITIZED VS. SYNTHESIZED SPEECH

Speech generation can be accomplished in several ways. Digitized speech is produced by converting analog speech signals to digital wave form. The computer records the waveform by sampling the signal's voltage periodically through an analog to digital (A/D) converter and then stores it as a binary value. The resulting binary data is then stored until needed at which time the original waveform is recreated by sequentially sending the stored values to a digital to analog converter (D/A) at the same rate as the original sampling.

There is a trade-off involved with digitizing speech. The bit density used to recreate the speech can be raised or lowered. Lowering the bit density obviously takes up less memory but the quality of speech is also degraded. Raising the bit rate improves the quality of the speech until it is nearly indistinguishable from analog recorded human speech but at the cost of a large amount of memory. Therefore, the user must decide on an appropriate compromise for a particular application.

Speech synthesis, another type of speech generation, typically employs a synthesis-by-rule scheme using formant-resonators. A formant resonator speech synthesizer models the human vocal tract and can reproduce the approximately 40 phonemes which comprise the English language. Phonemes may be defined as the set of the smallest units of speech that distinguish one utterance or word from another in a given language. High quality speech synthesis is dependent on how well transitions from one phoneme to another are handled, eg. from vowel to consonant and consonant to vowel. Furthermore, accuracy of the timing of the generated phonemic segments also contributes to the quality of the synthetic speech. Finally, the phonetic accuracy of the segments of speech are crucial to the production of high quality speech synthesis.

Text-to-speech rules, when used in conjunction with a speech synthesis technique, provide the user with real-time unlimited word production capabilities. Currently the text-to-speech software needed to produce unlimited speech generation capabilities requires approximately 16k of memory. Text-to-speech algorithms are a hierarchical set of linguistic rules and are entirely software based. When these rules are imposed on a particular synthesis technique, they provide the means whereby individual phonemes may be concatenated to produce realistic sounding speech.

The quality of speech synthesis when coupled with text-to-speech rules is dependent not only on how well the synthesis is executed but also on the particular linguistic rules which comprise the text-to-speech software. Since no standards pertaining to these rules have been created, they can be more or less accurate phonetically depending upon the manufacturer (Simpson, 1983).

In essence, the quality of synthesized speech is contingent upon both the hardware and software used to generate the speech. No one synthesis technique is intrinsically better than another. Rather, a particular technique's success or lack thereof is dependent upon how well it is executed (Simpson, 1983). Current speech synthesis technology tends to produce rather mechanical sounding speech. Listeners will often perceive a foreign accent in the speech produced by a synthesizer. This appears to be attributable to the fact that the rules that govern human speech code are very complex and the fact that not all of these rules are known at this time.

Today's speech generation technology, both digitization and synthesis, share a common weakness in determining the placing of articulation features for consonants. Further research is needed to determine exactly what speech cue makes us hear the place of articulation.

PERFORMANCE MEASUREMENT

Typically, intelligibility is used as the standard performance measure of both digitized and synthesized speech. There is a tendency, however, to measure intelligibility based on single words spoken in isolation, thereby eliminating any contextual cues that may aid in overall comprehensibility. Since human communications are rarely conducted in an isolated word fashion, a more realistic performance metric might be one in which intelligibility is measured for phrases, sentences, or some meaningful word group.

AUTOMATIC SPEECH RECOGNITION IN THE FLIGHT ENVIRONMENT

Although the pilot flying a high workload mission stands to gain tremendously from the use of voice command, the environmental, physical, and emotional factors impinging upon the pilot make speech recognition difficult to achieve reliably in the flight environment. Noise, vibration, stress, fatigue, and workload all act upon the pilot throughout any mission. These environmental and human effects manifest themselves to the speech recognition device as radically varying speech patterns for any given word in the operational vocabulary. Although problems such as noise and user stress and fatigue are not unique to the cockpit application of ASR technology, they are intensified and their effects are perhaps more critical than in industrial or office environments. However, the need to aid the pilot in his increasingly demanding job has motivated considerable research directed towards overcoming these problems. In the following section many of these factors will be examined.

AMBIENT NOISE

A major problem associated with the use of ASR in the flight environment concerns ambient cockpit noise and the creation of reference templates. Should an on board ASR system (either speaker dependent or independent) be trained in the presence of ambient cockpit noise, or will reference templates created in the presence of no noise be adequate for use in flight? Research conducted at NASA-Ames Research Center (Coler, Plummer &

Huff, 1983; Kersteen, 1982) indicates that when an isolated word ASR system is trained in a quiet environment and recognition is then attempted using these training templates in the presence of noise (95-100 dBA of helicopter noise), obtained recognition accuracy rates are quite low (78%). Conversely, if the system is trained in the presence of background noise and recognition is conducted in that same ambient noise level, accuracy rates are quite high (97%). These results are attributable to the fact that when training occurs in a relatively quiet environment and recognition then takes place in the presence of noise, the training templates simply do not reflect the noise component. Thus, the match between the templates and the incoming utterance is poor, yielding low levels of recognition accuracy.

Obviously, the need to create reference templates by iterating the entire operational vocabulary several times during flight is both distracting and annoying to the pilot. There are, however, several possible solutions. First, an algorithm that continually samples background noise and incorporates this noise into the reference template may alleviate the problem. However, there is currently no algorithm that can update the templates fast enough to keep up with rapidly changing ambient cockpit noise levels. Second, simulated cockpit noise may provide enough fidelity that a pilot could create adequate reference templates on the ground in the presence of this simulated noise. These templates would then be loaded into the aircraft avionics suite for use in flight along with other specifics. Finally, the use of better sound proofing materials in the cockpit may reduce noise to an acceptable operational level for an ASR device in future rotorcraft.

UPDATING REFERENCE TEMPLATES

A second major problem relates to the length of time one set of reference templates can be used before retraining is needed since speech patterns change with time, stress, and fatigue. Does the pilot need to train the ASR system prior to every flight or will one set of reference templates be valid for a week or a month given that the vocabulary does not change? Furthermore, will the pilot need to retrain the system on some words during the course of a mission? The effects of stress and fatigue on speech characteristics are more difficult to isolate because they can operate either singly or in combination on the pilot. Stress levels are likely to vary drastically during the course of a given mission. Does this mean that during times of high stress, incoming recognition utterances will be so different that accurate recognition can not occur? Once again, an algorithm that updates the reference templates not only with background noise characteristics but also with changing speech pattern characteristics may help solve this problem. Clearly, more research pertaining to the effects of time, stress, and fatigue on speech patterns is needed.

STORAGE MEDIA

A more technical issue related to the use of ASR in flight concerns the best storage media for the reference templates for the flight environment. A variety of storage devices are available, such as magnetic tape, bubble memory, etc. Furthermore, it is possible that magnetic strips like those found on credit cards may become available for the storage of reference templates. Whatever device is chosen for the cockpit application, it must be compact, lightweight, non-volatile, heat and shock resistant, and longlasting.

ACTIVATION OF THE VOICE SYSTEM

To use voice command in the cockpit, there must be some way to activate the speech recognition system. There are several alternatives for accomplishing this task; however, little or no research has addressed which alternative is the safest, most acceptable, and least obtrusive. One alternative is to install a push-to-talk switch in the cockpit. The pilot would have to activate this switch with each input to the recognizer. Another alternative would be to leave the device in a continual ready mode, with the hope that accidental activation does not occur. Finally, the device could be left in the ready mode, waiting for a key word which signals the device to prepare for input.

COMMAND LANGUAGE

It has already been mentioned that connected speech recognition capabilities are becoming commercially available. These capabilities will probably be expanded beyond the current ability to recognize connected digits by the mid 1990's timeframe. Connected word recognition capabilities (as opposed to isolated word recognition) are clearly needed in the cockpit if workload is to be reduced, rather than increased, with voice command. The nature of the command language and syntax structure used between human and aircraft deserves considerable attention. It is crucial that the command language be as natural for the pilot as possible. More specifically, pilots will accept and learn a language using "pilot jargon" more easily than an unnatural command language. Additionally, command sequences to an ASR device that capitalize upon the way a pilot normally interacts with another crewmember will be learned and remembered better. The naturalness of the command sequence will become critical during times of high workload when the pilot has little available capacity to remember a given command sequence. Furthermore, the command language and syntax structure must be flexible enough that the pilot can express a command to the ASR device in any of several ways. Again, this capability will reduce any additional cognitive burden associated with remembering a specific, rigid command sequence. In essence the command language used in a cockpit should be designed to reduce rather than increase the pilot's cognitive load.

RESEARCH ISSUES

Table 1 summarizes the research issues concerning the use of ASR in the helicopter cockpit

TABLE 1

Automatic Speech Recognition Research Issues

1. How should the degrading effects of background noise on ASR accuracy be dealt with in the cockpit?

2. How long can reference templates be stored and then used with acceptable recognition accuracy rates?

3. What effects do stress and fatigue have on speech patterns and hence on ASR accuracy?

4. If a reference template requires updating or retraining during flight, how should this be accomplished and how should the pilot be made aware of this requirement without disrupting primary flight tasks?

5. What storage media for the reference templates will be best for the flight environment?

6. What is the best way to activate the ASR device and prepare it for input?

7. If a connected word recognizer is used, how should the command language between the pilot and aircraft be structured?

SPEECH GENERATION IN THE FLIGHT ENVIRONMENT

Speech generation has been considered for two main functions in the cockpit: for conveying caution, warning and alert type messages and as a prompt or feedback response to voice recognition input. Voiced alert messages in the cockpit have been in existence for a number of years now. There are two advantages of this capability. First, it alerts or warns the pilot without diverting visual attention. Furthermore, voiced alerts or warnings convey more information than traditional bells, buzzers, tones, etc. Voice warnings have also been suggested for articulating system failures and threat detection messages in the LHX cockpit.

SYNTHESIZED VS DIGITIZED SPEECH

For the aircraft cockpit, synthesized speech is more flexible than digitized speech. Furthermore, a synthetic speech-by-rule system does not have the vocabulary limitations that are found in a digitized speech system. With digitized speech, every word needed for an application must be identified, digitized, and then stored. Synthesis systems have virtually unlimited vocabulary. Digitized speech systems pose two problems for an aircraft application: they limit flexibility in that the number of usable words is fixed, and vocabulary size must be kept at a minimum or memory requirements and access time becomes unacceptable.

By virtue of the fact that synthesized speech sounds mechanical, it works well as a voice warning system since it stands out against the background radio communications ongoing in the cockpit. Simpson (1980) purports that a high fidelity representation of human speech enunciating a warning message might very easily blend with other ongoing cockpit communications, whereas a more mechancial sounding speech will stand out.

INTELLIGIBILITY OF SYNTHESIZED SPEECH

An important consideration in the integration of speech synthesis in the cockpit relates to its intelligibility. Several researchers present evidence suggesting that rule generated synthetic speech may be less intelligible than natural speech or speech digitized at a high data rate. Using a MITalk unrestricted text-to-speech synthesizer, Pisoni and Hunnicutt (1980) found that phoneme recognition for synthetic speech was 93.1% compared to 99.4% for natural speech. These researchers concluded that the difficulties observed in the perception and comprehension of synthetic speech are due to increased processing demands in short-term memory.

An alternative explanation might be that the decrease in performance associated with synthetic speech is due to a lack of familiarity with its distinctive "accent". In other words, the intelligibility of synthetic speech might be no less than listening to a person speak with a foreign accent. The point to be made here is that there may be nothing inherent in synthetic speech that makes it less intelligible than natural speech. In fact it may be more accurate to regard the two as points on a continuum rather than as two separate entities. The intelligibility of human speech varies with the listener's familiarity with the accent as does the intelligibility of synthetic speech. Clearly, further research with respect to the issue of training and familiarity as it relates to the intelligibility of synthetic speech is needed prior to its integration in the cockpit.

A related issue is the need to compare the intelligibility and comprehensibility of various commercially available speech synthesis devices among themselves, rather than continue to compare human speech with one particular brand of speech synthesis. The comparison of human speech and synthesized speech has no point of reference if a baseline has not been established for the differential intelligibility of the various commercially available speech synthesis devices.

SPEECH PITCH AND RATE

In addition to the unlimited vocabulary capability provided by text-to-speech synthesis techniques, almost all speech synthesizers have adjustable speech pitch and rate capabilities. Though these additional capabilities provide flexibility to the user or system designer, their interactive and/or additive effects on intelligibility and comprehension need to be considered. Simpson and Marchionda-Frost (1983) conducted a study which addressed the effects of speech pitch and rate in the presence of 85 dBA of simulated helicopter noise. These experimenters hypothesized that synthesized speech with a fundamental frequency above the frequency range of the highest amplitude octave band of the background noise would be correctly perceived more often than speech with a fundamental frequency with the same octave band of background noise. This hypothesis was based on the assumption that background noise of the same fundamental frequency would mask certain perceptual features of the synthesized speech warning thereby causing a degradation in intelligibility. Although this hypothesis was not supported by the data, pitch of the synthesized speech warning should not be disregarded in further research. It is possible that the type of noise used (simulated helicopter noise) or the rather unrealistic loudness variability may have contributed to this variable's failure to reach significance.

With respect to speech rate, Simpson and Marchionda (1983) hypothesized that increasing the rate at which a message is presented (thereby decreasing the amount of time taken by the message itself) will reduce comprehension time. The elimination of redundant words from the message was also noted as a means of reducing the temporal length of the message. However, this method was disregarded since previous research suggests that this technique tends to decrease intelligibility and increase response time presumably because linguistic redundancy is an important perceptual feature of speech.

Interestingly, results indicate that increasing the speech rate to 178 words per minute (WPM) (maximum number of wpm tested) had no degrading effect on intelligibility and apparently reduced the time taken to comprehend the message. However, subjects (who were also pilots) indicated a preference for messages presented at a slightly slower rate of 156 wpm. At the fastest presentation rate (178 wpm) some subjects indicated that they feared missing parts of the message. The subjects also stated that the slow message rate (123 wpm) diverted their attention from the primary flight task because it took so long.

This research has a number of implications. The effects of synthesized voice pitch on intelligibility and comprehension deserves further research perhaps in a more realistic noise environment. The use of compressed speech has been suggested for use in the cockpit. Humans can process as many as 300 words per minute with sufficient training particularly if the information conveyed is expected by the listener and highly redundant. Voiced warnings and alerts in the cockpit are neither redundant nor expected. Furthermore, the pilot will be performing numerous other concurrent tasks while listening to voice warnings. It is likely that the use of compressed speech will increase rather than decrease the pilot's cognitive load. Furthermore, the temporal savings in reduced message length will probably not offset the cost in increased intelligibility. Conversely, synthesized voice messages presented at an unnaturally slow rate should be avoided in the cockpit since they appear to divert unnecessary amounts of attention.

INFLECTION RATE AND AMPLITUDE OF SYNTHESIZED SPEECH

Filtering techniques will soon become available with speech synthesizers that will allow the user to change the inflection rate and amplitude of the synthesized speech. This capability will permit a single speech synthesizer to produce different types of voices. The implication for the a cockpit application is the possibility of using different synthesized voices for different types of tasks in the cockpit. For example, changes in the amplitude of the synthesized voice warning could convey additional information as to the urgency of the warning ie. the louder the warning the more urgent. However, in implementing a display design such as this, the amplitude must be regulated so that the loudest warning does not overpower other cockpit communication. Conversely, the amplitude of the warning must not itself be overpowered by ambient cockpit noise. Clearly, additional research concerning the perceptual implications of these variables for a cockpit application is needed, particularly because they hold promise for enriching synthesized speech with more linguistic cues.

PRIORITIES OF VOICED MESSAGES, ALERTS, AND WARNINGS

Given that voice warnings are and will be used in the cockpit, a method must be adopted whereby these warnings can be assigned a priority in the event that several warnings need be conveyed simultaneously. On the assumption that one message can be presented at a time, the most important one must be relayed to the pilot first. Less important messages must be queued with respect to their urgency and then displayed following the pilot's acquisition of the most urgent message.

REPETITION OF VOICED INFORMATION

Related to the issue of setting priorities for voiced warning messages is the number of times a warning should be repeated to insure acquisition by the pilot. This issue can be approached in several ways; the message could repeat for a fixed interval of time, the pilot could turn it off, or the message could repeat until the problem was solved. In a study which specifically addressed cockpit voice warnings for air transport operations, Williams and Simpson (1976) reported that pilots prefer a cancel button to deactivate voice warnings at their discretion, especially if the warning is of high priority (demands immediate attention). Alternatively, a spoken command could also be used to end a warning. This study also revealed that pilots preferred to have other less critical warnings presented on a subsidiary display such as a CRT.

Not all of the messages presented to the pilot via speech synthesis will be of a mission-critical nature in the LHX. Speech displays may also be used to present information on request from the pilot. Regardless of the nature of the information , since speech is by nature temporally restricted, a visual replica of the auditory information should be provided to the pilot for later reference. In fact certain types of information could be presented to the pilot in hard copy format in conjunction with the auditory presentation. This approach is well suited to information needed which will be referred back to later by the pilot during the course of a mission. Specifically, weather and navigation information is well suited for hard-copy presentation.

RESEARCH ISSUES

Table 2 contains a summary of the research issues related to the use of speech synthesis in the cockpit.

TABLE 2

SPEECH SYNTHESIS RESEARCH ISSUES

1. What are the effects of training and familiarity on the intelligibility of synthesized speech?

2. How do the various commercially available speech synthesis devices compare with each other in comprehensibility and intelligibility?

3. How does the pitch of the synthesized speech effect intelligibility in the presence of actual helicopter noise?

4. What is the differential intelligibility and comprehensibility of different voice types provided by a single speech synthesis technique?

5. Is there an appreciable gain in information transmitted when several different voice types are used as opposed to just one?

6. Do several voice types complicate rather than simplify the pilot's task?

7. Do voice messages, alerts, and warnings need to be assigned priorities? If so, what is the optimum way to assign priorities?

8. How many times should a voiced warning be repeated?

9. How should voiced messages be terminated by the pilot?

10. Should there be a visual back-up display for an auditory display of information to the pilot?

11. Is there any information that should be presented to the pilot in hard copy (printout) format as opposed to soft copy (CRT) or auditory?

FUNCTIONAL DESCRIPTION OF THE LHX AVIONICS SUITE

The primary reason for the Army's development of the LHX family of light/scout attack helicopters has been the need for an all weather aircraft with day/night capabilities. The LHX also is being designed for defense of Army aviation. Mission requirements will demand a considerable amount of nap-of-the-earth (NOE) type flying, in which the helicopter is flying low and fast and avoiding obstacles. The most outstanding and challenging aspect of the LHX from a human factors design point of view is the Army's desire to limit the operation of this aircraft to a single crewmember. Current attack helicopter missions require both a pilot and co-pilot. The co-pilot, seated in front of the pilot, performs various weapon related functions and relays verbal navigation commands to the pilot whose primary task is manual control of the helicopter. Even with two crewmembers, workload is often quite high, especially during critical attack mission segments when simultaneous target detection and weapon release and control functions are occurring. Clearly, the development of a single pilot cockpit will rely heavily on higher levels of task automation than currently exist.

LHX mission functions can be generalized into four major roles for the pilot: flight, offense, defense, and mission management. Since the pilot can only fill one of these roles at a time, the other roles must be automated to avoid overloading him. This implies that the avionics system must allow the pilot to perform whatever task is primary at the moment and automatically perform the secondary tasks. These requirements are necessitating design of the LHX based on advanced technology, some of which may not yet be available. The avionics architecture will employ an array of sophisticated sensors and advanced concepts in integrating and controlling these sensors. The Army's desire for a one-man crew, coupled with the new and expanded mission capabilities, increases the need for innovative design of display and control modes for the pilot, as well as more automation.

As outlined in Honeywell's report to the Army Aviation Research and Development Command (conducted under DAAK50-81-C-0038) the primary subsystems which comprise the current LHX avionics suite are:

1) Navigation
2) Target Acquisition and Attack
3) Flight Control
4) Communication
5) Threat Defense
6) Data Management
7) Control and Display

The success of the LHX will depend upon the design of the control and display subsystem since this subsystem provides the pilot/aircraft interface. No amount of technology will make this aircraft fully operational unless a prior determination is made as to the type of information the pilot will need during various mission segments and the rate and sense modality in which this information should be transferred between the pilot and the aircraft. In an effort to facilitate this information transfer function between pilot and aircraft, the following concepts are being considered for integration into the LHX:

1) No windows. Due to the problems associated with infra-red radar signature, windows may be essentially eliminated from the LHX cockpit. Thus, a wide field of view (60 by 160 degrees) wrap around display will be used for pilotage and for the display of flight control, targeting, threat detection, and fire control symbology. This display will be consistent in terms of symbology among all conditions of day, night, and adverse weather.

2) A terrain mapping display. For further navigation functions, a digital terrain mapping display, operating from digital terrain data bases, will provide threat and battlefield information. Upon pilot request, this computer driven display will also have the ability to plot courses between known waypoints.

3) A "display-by-exception" concept. This will be used for system status monitoring in which information will be presented to the pilot only if it is mission critical. Unlike current cockpit design in which the pilot must scan numerous system status instruments continually during flight, the display-by-exception design will lessen the need for the traditional continuous instrument scan, thereby reducing visual workload.

4) Integrated and automated systems. These will be employed in an effort to minimize the number of frequently executed routine operations that a pilot typically performs.

5) Voice technology. Voice interaction with the various on board subsytems will be used in this aircraft in a further

attempt to reduce pilot workload so that one man operation is feasible. Automatic speech recognition (ASR) will be used as an alternate means of system control and for entering and receiving flight information. Speech generation will be used an an alternate means of information display.

Speech technology has been recommended specifically for the following functions in the LHX:

SPEECH RECOGNITION

1. Automatic target recognizer tasks
2. Sensor (selection, mode, lock-on)
3. Terrain map display (request updates)
4. System monitoring (request information)

SPEECH GENERATION

1. Alert and warning messages
2. Feedback

Speech technology was chosen for these tasks particularly to enhance performance in multiple-task situations where visual monitoring and manual control of critical tasks will be important.

SPEECH INTERACTION

A considerable amount of applied research has been directed towards the use of speech recognition (speech input) as an alternative to manual keyboard data entry and speech generation (speech output) as an alternative to the visual information presented on traditional aircraft anunciator panels. Optimal use of speech technology in the cockpit, however, will be in an interactive mode where speech input and output are logically combined. In designing a truly voice interactive system, attention must be given to easing pilot visual workload while avoiding pilot auditory overload.

Voorhees, Marchionda, and Atchison (1982) conducted a study in which they assessed the use of speech technology in a simulated helicopter NOE environment. Subjects in this study performed an extremely demanding visual/manual tracking task. Crucial airspeed, altitude, and torque information was presented to them in one of three ways. One group of subjects received this information by traditional panel-mounted instruments (thus requiring the subjects to divert attention from the primary task when they needed such information). Another group of subjects received the flight information in the form of thermometer-type gauges that were arranged on the periphery of the CRT on which the primary task was displayed. This condition simulated a head-up type display. In the third condition subjects received a visual display of only the primary task. When flight information was needed, the subjects asked for it in the form of a single spoken command, eg. "airspeed", "altitude", and "torque." After computer recognition of this command, synthesized speech feedback provided the necessary information to the subject. In this condition, the subject's visual attention could remain on the primary task at all times.

Results of this study indicated that flight performance in the voice interactive condition was significantly better than flight performance in the other two conditions. This study is interesting in that not only does it exemplify the merits, in terms of improved flight performance, of using the auditory/vocal channels as a means of acquiring information in a demanding flight task. It also suggests that although HUDs eliminate the need for the pilot to scan an instrument panel, there still may be some unwanted diversion of visual attention associated with the use of these displays.

As mentioned earlier, a number of voice tasks have been recommended for integration in the LHX. One particular subset of LHX functions may involve both speech input and output in the use of an automatic target recognizer (ATR). In an ongoing effort to develop an ATR for LHX attack and scout missions, Honeywell has designed a Prototype Automatic Target Screener (PATS). This system is capable of sensing, identifying, and classifying ground targets using forward looking infra-red (FLIR) or day TV imagery. In conjunction with the development of PATS, Mountford, Schwartz, and Graffunder (1983) identified the following pilot interactions with PATS that lend themselves to speech technology implementation:

1. Enter navigation coordinates for recognizer search area

2. Select modes of PATS operation: search and designate

3. Request display of another detected target

4. Modify detection confidence criteria

5. Change target priorities

6. Assign weapons to targets

7. Retrain/reinforce target indentification algorithm

Mountford, Schwartz, and Graffunder (1983) created a simulation in which several of the PATS tasks (1,2,3, and 6) were combined with a concurrent tracking task. The navigation-targeting-weapon selection sequence of tasks associated with PATS was performed repeatedly according to the following three task control, interaction, and feedback formats:

	Input Modality	Feedback Modality
1.	Manual	Visual
2.	Speech	Visual
3.	Speech	Speech

The overall results of this study indicated a dual-task (PATS and tracking tasks) performance advantage for speech-speech data input as opposed to manual-visual data entry. Although tracking performance error doubled when the tracking and PATS tasks were performed concurrently, tracking error was lower when speech input and output were used interactively than when speech-visual or manual-visual input and output modalities were used for the PATS task. Mountford et al. attribute the performance advantage for speech input and output to the freeing of visual and manual resources so they can be dedicated solely to the tracking task.

Results of this study also indicated that, particularly for the navigation tasks, the time to complete this task was greatest in the speech-speech modality. This result is not surprising since the navigation task required the input of strings of digits with feedback for each one. The time handicap for navigation digit entry using speech could be overcome by the use of a connected speech recognizer, which would allow the pilot to string the digits together as one data entry as opposed to several discrete entries. However, since speech is temporal by nature, the additional time needed for the articulation of feedback messages is inherent to this mode of information transmission.

Thus, it appears that there is experimental evidence suggesting that speech is desirable for the acquisition of information in a demanding flight task. The next question is how should this voice interactive dialogue between human and machine be designed? Either speech or manual input to the avionics suite requires verification that the correct input was received. In a non-critical mission segment, visual feedback supplied via CRT may be adequate. However, during mission segments in which heavy visual demands are placed upon the pilot, auditory feedback will be most desirable, as will voice input. Taken a step further, structuring the interactive dialogue between the pilot and the aircraft will be facilitated with the additional capabilities proffered by speech input and output. Currently, information has been presented visually to the pilot and controlled through multifunction keyboards. The addition of speech I/O to future cockpits will provide complete hands-off, eyes-off interaction with various on board systems.

In structuring this dialogue it is critical that the user be supplied with the knowledge that previous responses have been input and recognized by the system correctly. Mountford, North, Metz, and Warner (1982) have examined three types of dialogues for man/machine communication which they characterize as "succinct", "intermediate", and "verbose" depending upon the wordiness of the dialogue. Results of this study indicate that succinct dialogues are preferable to the more verbose dialogues primarily because they require less involvement from the pilot in terms of time and attention. This work highlights the importance of keeping aircraft/pilot interactions brief and to the point.

Furthermore, interaction between pilot and aircraft systems must be as natural for the human as possible. One of the advantages of using speech as a mode of interaction with on board systems (as opposed to visual/manual interaction) is that speech is the most natural mode of communication for humans. Efforts must be made to capitalize on this naturalness by incorporating enough flexibility into this communication link so the pilot can communicate his/her intentions to the aircraft in much the same way as she/he would to another crewmember. Conceptualizing and creating an optimal voice interactive dialogue based on pilot-to- co-pilot communications will necessarily require considerable thought and artificial intelligence. In human communication, specifically pilot/co-pilot communications, a great deal of intent is inferred by the crewmembers involved in the communication. This means that certain things are done or assumed by the communicators based on the characteristics of the given situation. This implies that somehow the machine must be apprised of or be made smart enough to infer certain mission and situation specifics. The accomplishment of this non-trivial task will provide the added flexibility characteristic of human communication to the man/machine communication link that will begin to allow full realization of the potential for speech technology in the cockpit. The purpose of this paper is not to expound upon artificial intelligence and its many cryptic interpretations. Let it suffice to say that heightened and continuing research in this area will be highly beneficial to the creation of this very important communication link in future generation rotorcraft.

An issue that is presently under debate relates to whether the pilot should be provided with reversionary controls in the event of a voice system failure. Should there be a manual backup for tasks that have been allocated to voice command; and should there be visual backups for auditory displays of information? Reversionary controls may be important for psychological as well as technical reasons. Certain situations may arise in which the pilot will simply feel more comfortable performing a task manually rather than verbally.

RESEARCH ISSUES

Table 3 contains a summary of the research issues related to speech interaction in the helicopter cockpit.

TABLE 3

SPEECH INTERACTION RESEARCH ISSUES

1. With respect to the structure of a voice interactive system between pilot and aircraft (avionics suite), it has already been established that succinct dialogues are preferable to wordy dialogues. What other general rules can be derived to govern the integration of speech interaction in the cockpit?

2. Does the pilot need reversionary controls?

3. What are the psychological implications of not providing the pilot with reversionary controls?

AUTOMATIC SPEECH RECOGNITION TECHNOLOGY ASSESSMENT

The specifications outlined in this technology assessment are slanted towards those which are of importance in a cockpit application. The specifications are by no means exhaustive and may not do justice to some of the products that have been developed for applications other than those involving cockpit integration.

An issue which needs clarification prior to reading this assessment is the configuration of the various speech recognition products. There are three basic types of configurations into which most speech products fall. First, the technology may be integrated into a "development system." This means that it has been factory interfaced with a computer prior to its purchase by the user. "Development systems" typically come with software to aid in the application development. Second, there are "standalone" systems that communicate with the host processor chosen by the user. This means that the user buys a board-level product and then interfaces it to his or her own host computer. Typically, this type of system requires the creation of a considerable amount of software on the part of the user. Finally, speech recognition products may be in the form of standard or custom OEM chips to accommodate a wide range of form factor and interface requirements.

In many cases, this technology assessment details only one product from a particular manufacturer. This does not mean that manufacturer does not have numerous speech products available, it simply means that the system chosen for assessment is the one most feasible for a cockpit application.

Another issue that must be clarified before continuing with the assessment concerns flightworthy ASR systems. Several manufacturers are currently working on these; however, it must be noted that these systems are still in the design and developement phase, with a considerable amount of work still needed to make their use feasible in the flight environment.

First, a table (Table 4) will be presented in which the pertinent specifications for each speech recognition device are summarized. This will be followed by a more detailed description of each of the assessed devices.

AUTOMATIC SPEECH RECOGNITION HARDWARE SPECIFICATIONS

SPECIFICATION	INTEL	INTERSTATE SYS 300	ITT	LEAR-SIEGLER	NEC DP 200	NEC SR 100	SCOTT VET 2	VERBEX	VOTAN V5000
SPEAKER INDEPENDENT/DEPENDENT	DEPENDENT	DEPENDENT	DEPENDENT	DEPENDENT	DEPENDENT	DEPENDENT	DEPENDENT	DEPENDENT	DEPENDENT
ISOLATED WORD RECOGNITION?	YES	YES	YES	YES	YES	YES	YES	YES	YES
CONNECTED/CONTINUOUS RECOGNITION?	NEITHER	NEITHER	NO	NO	CONNECTED	QUIKTALK*	NEITHER	CONTINUOUS	NO
FORM OF HARDWARE	BOARD*	BOARD	FLIGHTWORTHY SYSTEM	FLIGHTWORTHY SYSTEM	2 BOX SYSTEM	BOARD	BOARD AND SOFTWARE	SPECIAL SYSTEM	BOARD
NUMBER OF TRAINING PASSES	2*	2-5 PASSES	UNSPECIFIED	3 PASSES	1-2 PASSES	1-2 PASSES	3-5 PASSES	NONE	1-2 PASSES
MAXIMUM VOCABULARY SIZE	200 WDS	100 WORDS AND/OR PHRASES	200 WD	100 WDS	150 WDS. CONNECTED 500 WDS. DISCRETE	120 WDS.	40-89 WDS	120, 240 or 360 WDS	80-255 WDS
RECOGNITION RESPONSE TIME	LESS THAN 500 MS	(50 + N) MS	LESS THAN 250 MS	LESS THAN 400 MS	300 MS AT END OF SENTENCE	300 MS (DISCRETE MODE)	250 MS	VARIABLE	180 MS*
SELECTABLE REJECTION THRESHOLD	YES	YES	YES	YES	YES	YES	NO	N/A	YES
MINIMUM PAUSE BETWEEN UTTERANCES	50 MS	160 MS	?	100 MSEC	N/A	30 MS	250 MS	CONTINUOUS	180 MS*
MINIMUM WORD LENGTH	*	80 MS	UNSPECIFIED	NO	.2 SEC PER SENTENCE	.3 SEC	120 MS	N/A	NONE
MAXIMUM WORD LENGTH	*	1.25 S	NO	3 SEC	4 SEC PER SENTENCE	2 SEC (4 SEC QUIKTALK)	1.5 SEC	N/A	1.5-2.0 SEC
POWER REQUIREMENTS	+5V, +12V	+5V @ 1 AMPERE; +12V @ 100 MILLIAMPS	115V AC, 400 HZ	115V AC, 400 HZ; 28V DC	90V - 127 V AC, 50/60 HZ.	115V AC ± 10%, 50/60 HZ	115V AC, 60 HZ	115V AC, 60 HZ	115V, 50/60
WEIGHT	?	.5 KG	APPROX 19 K	DEVELOPMENT	33.4 KG (70 LBS.)	3.5 KG	APPROX 2 KG	227 KG	*
DIMENSIONS	178 MM X 127 MM X 305 MM	432 MM X 89 MM X 318 MM		DEVELOPMENT	N/A	240 MM X 350 MM X 60 MM	44 KG X 203 MM X 279 MM	560 MM X 840 MM X 180 MM	279 MM X 457 MM X 89 MM
INTERFACE REQUIREMENT	MULTIBUS, RS 232C	RS 232C	MIL-STD 1553B, RS 232C	MIL-STD-1553B, 20mA CURRENT LOOP	RS 232C EIA RS422, IEEE 488	RS 232C, CCITT V.24 + V.28	APPLE	RS 232C, 20 MA CURRENT LOOP	RS 232C
RETRAIN CAPABILITY	YES	YES	YES	YES	YES	YES	NO	NO	YES
BUILT-IN SELF-TEST	YES	YES	YES	YES	YES	YES	NO	YES	NO
TEMPLATES AVERAGED/ MAINTAINED SEPARATE	AVERAGED*	AVERAGED	AVERAGE	BOTH	N/A	SEPARATE	AVERAGED	N/A	SEPARATE
PRICE	$2900.00-BOARD $4900.00-SOFTWARE*	$1995.00	DEVELOPMENT	DEVELOPMENT	$15,000.00	$2,000.00	$795.00	$80,000.00	$5000.00
SPEECH OUTPUT CAPABILITIES	NO	NO	YES*	YES	YES	YES	NO	YES	YES DIGITIZED
CONSTRUCTED FOR COCKPIT APPLICATION	NO	NO	YES	YES	NO	NO*	NO	NO	NO*

* SEE WRITTEN DESCRIPTION FOR MORE DETAIL

Intel

Intel produces what they call a speech transaction family of speech products. The speech transaction board is available for \$2,900.00 and is the actual speech recognition hardware. The speech transaction development set (\$4,900.00) is the accompanying operating system and software which allows the user to integrate the hardware into an actual application. Intel is currently in the process of making several major updates to their speech transaction family of products; an improved recognition algorithm which will provide better constant discrimination will be implemented for the speech transaction board. In addition, the ability to maintain several templates for each vocabulary word may also be implemented. For this reason, the number of training passes needed to use this device is undecided. Additional noise processing will be added with an algorithm that will measure the background noise between words and subsequently subtract this noise from the speech signal. The impulse noise filter will also be enhanced. The actual levels of noise to which this device is immune are as yet undetermined. The speech transaction development set will also be expanded with additional software.

Although Intel does not specifically offer speech output capabilities with this system, they have provided the means for the user to integrate his or her own speech synthesis device with this system. Intel anticipates a full release of these expanded capabilities in November, 1983.

Interstate

Interstate offers a wide variety of speech recognition equipment, ranging from chips to fully integrated voice recognition terminals. Many of their products are designed to operate with specific host computers such as Lear Siegler Incorporated and Digital Equipment Corporation computers. Interstate also offers several types of speaker-independent speech recognition chips.

SYS 300

The SYS 300 is a board-level, speaker-dependent speech recognition system designed specifically to be interfaced to most RS 232C terminals. There are approximately 15 recognition commands that may be used in creating application software for this device. The device is capable of recognizing up to 100 words. Interstate claims that the SYS 300 is resistant to noise levels up to 80 dB(A). A voice output module (VTM 150) may be purchased for $995.00 and interfaced to the SYS 300. The VTM 150 includes a 500-word fixed vocabulary and 1000 word user-programmable vocabulary with text-to-speech capabilities.

ITT

ITT has developed a flightworthy speaker-dependent isolated word recognition system for the tactical aircraft cockpit environment. This system was designed to withstand the high "g" levels, high noise levels, and oxygen mask breath noise inherent in the tactical aircraft cockpit.

To a large extent, this device is still in the developmental stage. However, preliminary flight testing aboard the Air Force Technology Integrator (AFTI) F-16 indicated that the ITT system maintained a recognition accuracy rate of approximately 90% in high "g" and noise levels (5 "g" and 115 dB(A), respectively). ITT is working to integrate speech synthesis capabilities in this device as well as connected word recognition capabilities.

Lear Siegler Incorporated

Lear Siegler has also developed a flightworthy, tactical Voice Controlled Interactive Device (VCID) for military application flight testing. This system was designed to operate in the same operating environment as the ITT speech recognition device. Lear Siegler claims that this system can be trained on the ground in a low noise environment prior to use in flight. This device can accomodate a maximum vocabulary size of 256 words or short phrases. A speech synthesis unit will be available with the VCID for operator feedback.

The VCID has undergone preliminary flight testing aboard the AFTI F-16. Results indicated that in noise levels up to approximately 103 dB(A), recognition accuracy rates were in the 80% to 90% range. Beyond 103 dB(A), however, recognition accuracy declined abruptly. During later portions of these flight tests, Lear Siegler added a Speech Enhancement Unit (SEU) to the VCID which appeared to raise these recognition accuracy rates by several percentage points. The SEU is basically a front-end processor which samples the background noise and subtracts it from the speech signal.

Lear Siegler is currently making modifications to the VCID in preparation for the second phase of flight testing aboard the AFTI F-16. One of these modifications may be the ability to maintain separate templates for each word in the vocabulary. Lear Siegler is also working on connected word recognition capabilities for integration into the VCID.

NEC

NEC has two speech recognition system on the market: the SR 100 and the DP 200.

DP 200

For $15000.00 NEC offers a speech recognition system that will provide the user with up to .20 s of connected word recognition. A maximum vocabulary of 150 words can be used in the connected mode, and a maximum of 500 words can be recognized in the discrete mode. The DP 200 comes with two floppy disc drives, an operating system, and various software. Template handling can be done either internally in the DP 200 or through the host computer. The benefit associated with allowing the templates to be handled by the DP 200 is that it frees the host from continually having to monitor the interface line. This internal control process essentially preprocesses and buffers the incoming speech information before sending it to the host. The DP 200 requires minimal training and provides a retrain capability for select parts of the vocabulary. NEC claims that the DP 200 will withstand up to 85 dB(A) of random noise.

Speech output capabilities may be added to this system for an additional $4,600.00. This audio response unit uses a digitization technique and will provide the user with either 90 s of speech at 16 kb or 60 s of speech at 32 kb.

SR 100

The SR 100 is the only low cost ($2000.00) speech recognition product on the market that has connected word recognition capabilities. This high speech option or Quiktalk mode allows a maximum string of 10 words to be recognized in a connected fashion. Two training passes are required for these 10 words. In both the discrete and Quiktalk mode, the SR 100 maintains each template separately in memory.

To interface the SR 100 to a host computer, there are seven user definable parameters. For an additional $2000.00 a voice output device (AR 100) can be purchased to work with the SR 100. The AR 100 provides 120 s of digitized speech.

Although the SR 100 was not designed for use in the aircraft environment, United Technologies conducted an in-house test of the SR 100 in three noise conditions using two different types of microphones. The noise levels tested were 20 dB ambient noise, 85 dB S-76 cockpit noise, and 100 dB UH-60 cockpit noise. The tests were conducted using both a throat microphone and a Shure noise cancelling microphone. For the digit vocabulary using the throat microphone, the SR 100 achieved a recognition accuracy rate of 96% across all three noise conditions. Using the Shure noise cancelling microphone, an accuracy rate of appoximately 97% on the digit vocabulary was obtained across all three noise conditions.

<u>Scott Instruments</u>

Scott Instruments offers a low cost ($795.00) speaker dependent speech recognition system . The Voice Entry Terminal (VET) includes a terminal, a microphone, a microcomputer interface, user's manual, and system software. The VET is designed specifically to work with Apple Computers. The noise immunity of this system is unspecified.

Verbex

The Verbex 1800 is a high cost ($80,000.00) speaker-independent, connected word recognizer. This device was designed specifically to allow a user to communicate with a computer or a telephone switching system by talking to it over any telephone. The system can accomodate up to eight users simultaneously. Speech output (digitized speech) is an option with the Verbex 1800. The minimum recognition vocabulary consists of 10 digits (0-9) and "yes" and "no". This vocabulary may be expanded up to 50 words. The speech output vocabulary includes up to 32 words or 16 s of speech and can be expanded up to 512 words or 256 s of speech.

Votan

Votan offers the following types of speech technolgy: speaker-dependent and independent recognition, speech output, voice store and forward, vocoding, and speaker verification. Various combinations of these features are available either in system, standalone, or board form.

V5000

The V5000 combines speaker-dependent word recognition, speech output, and voice store and forward capabilities in a standalone unit. Votan's speech recognition technology requires one or two training passes and the resultant templates are stored separately in memory. The recognition response time for the V5000 is 180 ms plus an additional 2 ms for each word in the vocabulary. It must be noted that if syntax structures are used, the response time would be 180 ms plus 2 ms for each word in the syntax node (as opposed to 2 ms for each word in the entire vocabulary).

The speech output available from Votan is digitized and is user programmable. The user has a choice of three bit rates for the speech digitization.

The voice store and forward technology allows speech to be digitized, compressed, and stored in RAM memory. The speech can then be transferred to a host processor or a mass storage device. This information may be retrieved in audio form by reconverting the digital data back to an analog signal.

The noise immunity of the V5000 was recently tested at NASA-Ames Research Center (Coler, 1982). For the purposes of this test, the V5000 was trained on the digit vocabulary (0-9) in quiet and recognition was attempted both in quiet and in 100 dB(A) noise. The V5000 was also trained in 100 dB(A) noise and recognition was attempted again in both quiet and 100 dB(A) helicopter noise. Results indicated that from a grand total of 3,200 utterances (collected from eight subjects) only one miss or substitution error occurred and there were no rejections.

Votan is currently working on making continuous word recognition capabilities available as a firmware update to the V5000 by early 1984.

REFERENCES

Coler, C.R. "Helicopter speech command systems: Recent noise tests are encouraging". Speech Technololy 1, (3), p. 76-78, Oct. 1982.

Coler, C.R., Plummer, R.P. & Huff, E.M. NASA-Ames Research Center, Mountain View, CA, 1983 (Manuscript in preparation).

Kersteen, Z.A. "An evaluation of automatic speech recognition under three ambient noise levels". Proceedings of the Workshop on Standardization for Speech I/O Technology. National Bureau of Standards, Gaithersburg, MD, 1982.

Lea, W.A. Computer Speech Recognition: Seminar Workshop for a Short Course. Santa Barbara, CA: Speech Science Publications, 1979.

Mountford, S.J., North, R.A., Metz, S.V. & Warner, N. "Methodology for exploring voice-interactive tasks: Optimizing voice interactive dialogues". Proceedings of the Human Factors Society 26th Annual Meeting. Seattle, WA, 1982.

Mountford, S.J., Schwartz, J. & Graffunder, K. "Voice interactive prototype automatic target screener", Contract No. NAS2-11348, Honeywell Systems and Research Center, Minneapolis, Minn., April 1983.

Pisoni, D.B. & Hunnicutt, S. "Perceptual evaluation of MITalk: The MIT unrestricted text-to-speech system". Proceedings of the 1980 International Conference Record in Acoustics, Speech and Signal Processing. New York: IEEE, 1980.

Poock, G.K. "Voice recognition boosts command terminal throughput", Speech Technology 1, (2), p. 36-39, April 1982.

Simpson, C.A. Personal Communication, 1983.

Simpson, C.A. "Synthesized voice approach callouts for air transport operations", Report No. NASA CR-3300, Psycho-Linguistic Research Associates, 2055 Sterling Ave., Menlo Park, CA, 1980.

Simpson, C.A. & Marchionda-Frost, K. "Synthesized speech rate and pitch effects on intelligibility of warning messages for pilots". Proceedings of the Second Symposium on Aviation Psychology, Ohio State University, Columbus, OH, 1983 (In press).

Williams, D.H. & Simpson, C.A. "A systematic approach to advanced cockpit warning systems for air transport operations: Line pilots preferences", Proceedings of the Aircraft Safety and Operating Problems Conference. NASA Langley Research Center, Hampton, VA, Oct 1976.

Voorhees, J.W., Marchionda, K.M. & Atchison, V.L. "Auditory display of helicopter cockpit information", Proceedings of the Workshop on Standardization for Speech I/O Technology. National Bureau of Standards, Gaithersburg, MD, 1982.

APPENDIX

PILOT QUESTIONNAIRE

THE ROLE OF SPEECH INPUT AND OUTPUT IN THE HELICOPTER COCKPIT

At NASA-Ames Research Center, we are currently examining the potential uses for voice warning and control systems in future helicopter cockpits. As you know, current rotorcraft operations require manual input (in the form of switch manipulations, flight control, etc.) to the various on board systems and provide visual and auditory output to the pilot in the form of flight instrument displays and alerting signals (horns, buzzers, etc.). We have acknowledged that the visual and manual demands placed on the helicopter pilot are at times excessive. Our work on speech technology in the cockpit is aimed at reducing or offloading these demands as well as increasing the utility of the aircraft.

An avionics system into which speech technology is integrated would involve "speaking" to an on board computer commanding it to perform switch sequences, requesting information from the various aircraft systems, etc. The system would recognize your voiced command, perform the requested task, and report back verbally, if requested, that the task has been completed. In addition the system could give you warning and advisory information verbally rather than visually.

This questionnaire is divided into two sections. In the first section, we have listed some tasks that might be performed by voice in an existing helicopter, the AH-1. Because you have had experience flying this helicoper, we would like you to evaluate each of these tasks with respect to the potential desirability of having speech perform these tasks. When you respond to these questions assume that a computer has been added to the aircraft and that you have the ability to control on board systems by voice and receive various types of information verbally from this system. Our goal in this part of the questionnaire is to determine what types of tasks you think will be best suited for voice technology.

The second section of the questionnaire will give you the opportunity to think about the design of future rotorcraft and to tell us what you would like if virtually any cockpit design becomes possible.

QUESTIONNAIRE

The following questions are to provide you with an opportunity to contribute your ideas and opinions about how speech technology might be implemented in the AH-1. Your ideas are valuable and important since the information obtained from this questionnaire will provide guidelines for the implementation of this technology in future rotorcraft.

The personal data sheet is for the purpose of data analysis only. No comments or answers will be associated with your name.

Please answer each question carefully. The more comments and examples you have with respect to these tasks the better (please write them on the back of the page).

The five point scale provided after each question provides a continuum of desirability, from extremely undesirable to extremely desirable. Please indicate your opinion by circling the number which best describes your opinion.

EXAMPLE

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

PERSONAL DATA

Name/Rank____

Organization__________________ Position__________________

Age________ Date______________

PILOT EXPERIENCE. Please approximate hours by type.

Rotorcraft type	Hours Total
________________________	________________
________________________	________________
________________________	________________
________________________	________________
________________________	________________

Do you or have you flown fixed wing aircraft?

Yes____________ No______________

Aircraft type	Hours Total
________________________	________________
________________________	________________
________________________	________________
________________________	________________

Do you play video games?

Often________ Occasionally________ Never________

Do you own a home computer?

Yes___________ No_____________

Have you taken any computer programming courses?

Yes______________ No________________

Have you ever written a computer program?

Yes______________ No________________

Have you ever heard computer generated speech?

Yes______________ No________________

If yes, please explain.

____

____

____

Have you ever used an automatic speech recognition device?

Yes______________ No________________

If yes, please explain

____

____

____

Please provide other comments on attitudes, education or experience that might influence your answers to this questionnaire.

____

____

____

____

PILOT QUESTIONNAIRE

Computer generated speech could be used to advise you when certain parameters or systems move outside safe operating limits or become inoperative. A number of these are listed below. For each one rate how desirable it would be to have an advisory or warning about it presented by voice.

1. ENGINE OIL TEMPERATURE

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

2. ROTOR RPM

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

3. ENGINE RPM

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

4. TORQUE PRESSURE

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

5. TGT (Turbine Gas Temperature)

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

6. ENGINE OIL PRESSURE

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

7. ENGINE OIL BYPASS

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

8. FWD or AFT FUEL BOOST

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

9. ENG FUEL PUMP

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

10. 10% FUEL REMAINING

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

11. FUEL FILTER

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

12. XMSN OIL BYPASS

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

13. XSMN OIL PRESS

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

14. XSMN OIL HOT

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

15. HYD PRESS #1 or #2

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

16. INST INVERTER

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

17. DC GENERATOR

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

18. CHIP DETECTOR

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

19. IFF

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

Given the installation of a variety of sensors throughout the aircraft, advisory and/or warning information could be presented to you by voice. Items 20-23 deal with this type of information. For each one please indicate how desirable it would be to have this information presented to you by voice.

20. Advise that the helicopter has not been grounded during refueling or when it is being parked.

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

21. Advise if ground safety pins have not been installed in the pilot and/or gunner canopy removal arming/firing mechanisms when the helicopter is to be parked.

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

22. Warn if carbon monoxide, smoke etc. is detected in the cockpit.

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

23. Advise if stores jettison safety pins have not been installed when helicopter is on the ground.

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

24. How desirable would it be to have a voice generator assist in performing checklist items?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

25. Would it be desirable to have exact threat information details presented to you verbally. Eg. "SA10, 4 O'clock, launch?"

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

26. Assume for the moment that your aircraft is data-linked to the ground, would you find it desirable to be able to request targetting information and receive it verbally?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

27. Assume that the entire aircraft manual is stored in the aircraft computer's memory and is accessible to you during flight. Would it be desirable to request information from the manual by voice command and receive it verbally?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

28. Would you like to be reminded when certain tasks should be done, for example, "change IFF".

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

29. Can you think of any other type of information you would like to receive from a voice generator?

____

____

____

30. Further Comments:

31. Please rank the following three ways in which warning information could be presented to you by voice. A rank order of one (1) means most desirable and three (3) means least desirable. Please use each ranking only once.

_________The voice generator would say something like "Caution", which would then alert you to look to the instrument panel for a problem.

_________The voice generator could tell you exactly what is out of tolerance, eg. "Warning, oil pressure low".

_________The voice generator could tell you exactly what is out of tolerance, by how much, and a recommended course of action.

Some other method. Please elaborate.

32. If a voice generator is used as an aid in performing checklist items, please rank the following ways in which it could be implemented. A rank order of one (1) means most desirable and four (4) means least desirable. Please use each ranking only once.

________The voice generator could call out each item in the checklist for you to perform.

________You could run through the checklist, following which the voice system could remind you of any items that may have been overlooked or for any conditions which might preclude safe operations.

_________All checklist items could be placed under computer control and performed automatically for you. The voice generator would then advise you when the checklist had been completed or if any irregularities were encountered.

__________The voice generation system could call out the next item in the checklist when you request it to.

33. The following items comprise the general categories of functions for which computer generated speech might be used. Please rank these categories from one to four, with one (1) meaning the most desirable for computer generated speech and four (4) meaning the least desirable. Please use each ranking only once.

__________Presentation of advisory and cautionary type information eg. "oil pressure low".

__________Presentation of general information that has been requested by the pilot eg. "EGT 670 degrees".

__________Presentation of feedback or acknowledgment that tasks have been completed, for example, "Outboard stores selected".

__________Presentation of emerging information eg. "rotor RPM low".

A voiced command could be used to access information from various flight instruments. By saying, for example, "Request altitude", the system would come back with "125 feet". For each of the following six types of information, please rate how desirable it would be to request this information with a spoken command. When responding to these items, assume that the machine will recognize your voiced command with the accuracy of a human listener.

34. AIRSPEED

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

35. ALTITUDE

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

36. VERTICAL VELOCITY

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

37. OAT (Outside Air Temperature)

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

38. HEADING

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

39. TIME

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

40. TORQUE

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

41. ROTOR RPM

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

42. How desirable would it be to turn cockpit lighting on and off by voice command?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

43. Voice command could be used to reset circuit breakers. How desirable would it be to perform this task by voice?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

44. How desirable would it be to tune the radios by voice?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

45. If you had to use voice command to tune radios, would you rather tune the radio by frequency (eg. "Tune 256.4") or by name of the station (eg. "Tune Moffett Tower"). Please circle your preference.

FREQUENCY NAME

46. Would you find it desirable to configure the voice security equipment by voice, using one command to accomplish all the tasks. For example, "Set plain mode."

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable.

47. Similarly, would you like to configure the ADF by saying "Tune Evansville NDB, loop mode"

The computer would then perform the following tasks for you:

A) Tunes Evansville ADF
B) Identifies the station
C) Indicates whether you are in receiving a reliable signal from the station

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

48. A voice command could be used to request fuel required and burn-out times during a mission. How desirable would it be to perform this task by voice?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

49. A voice command could be used to select the type of weapon you wish to use in the AH-1, by saying, for example, "Select turret". Would this be a desirable candidate task for speech input and output.

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

50. Voice command could be used to select the particular weapon station you wish to use. Would it be desirable to perform this task with speech?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

51. The number of weapons to be fired could be specified by voice command. Would this be a desirable task for speech?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

52. The firing sequence of these weapons could also be specified by voice command. Would it be desirable to perform this task by voice?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

53. Voice command could be used to fire weapons. Would it be desirable to have this capability?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

54. Would it be desirable to jettison stores when necessary by voice command?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

55. If your aircraft was equipped with an automatic hover hold and bob-up mode, how desirable would it be for you to control these modes by voice command?

1	2	3	4	5
Extremely Undesirable	Somewhat Undesirable	Not Sure	Somewhat Desirable	Extremely Desirable

56. The following items comprise some of the general categories of tasks for which voice command might be used in helicopter operations. Please rank the desirability of performing these types of tasks by spoken command from one (1) to seven (7), with one meaning the most desirable and seven meaning the least desirable. Please read all areas before ranking them, and use each ranking only once.

__________ Vehicle control, for example, "Bob-up"

__________ Weapon stores management, for example, "select outboard stores".

__________ Navigation tasks, for example, "Tune Evansville VOR"

__________ Communications, eg. "Tune Moffett Tower"

__________ Subsystems management, for example, "HUD on"

__________ Weapon delivery, eg. "Launch TOW"

__________ Requesting flight instrument information by voice command and receiving that information from a speech system, eg. "Torque" - "88 Percent".

57. If you had the ability to use voice command in the cockpit, there are several ways in which you could activate the system (ie. let it know that your are talking to it). Please rank the desirability of the following activation methods from one (1) to three (3) with one meaning the most desirable and three meaning the least desirable.

__________ Push-to-talk switch

__________ Have the voice system actively listening for your spoken command all the time.

__________ Say a keyword which would activate the system prior to speaking the actual command.

Some other method. Please elaborate.

____

____

____

58. Comments: Please comment on the use of voice command for tasks in the above categories. Give us examples of any other categories of tasks for which voice command might be used in the AH-1.

59. Any other comments, general or specific on speech input and output applications in rotorcraft operations.

Go to next page

Section 2

In the following section we would like you to think seriously about how you would design a future pilot/aircraft interface (cockpit) in an effort to make your job easier and safer. Don't worry about whether your ideas are technologically feasible--treat them as if anything is possible. You will be given six areas to respond to; answer them from a scout/attack type of mission standpoint. Within each area, tell us what cockpit changes you would like to see in current rotorcraft, what you would like in a future rotorcraft, and how you would like to have it done. If you discuss a design change in an existing rotorcraft, be sure to specify which one. We also want to know how you would like to interact with your helicopter in each of these areas. In other words, for each change or idea you have, tell us whether you would like to use speech input and/or output, visual/ manual input and output, some combination thereof, or something completely different. Sketches, if applicable, might help us understand your ideas better.

Things to remember when completing this section

1. Be specific
2. Don't worry about writing style, etc. (just be legible)
3. Disregard current technological constraints
4. Sketch your ideas on the back of each page if you like.

If anything were possible, how would you improve and/or redesign each of the following areas of helicopter operations.

1. Navigation

2. Target acquisition and attack

3. Flight Control

4. Communications

5. Threat Defense

6. Mission Management

Thank you for your time and thoughts.