

THE ROLE OF VOICE TECHNOLOGY IN ADVANCED HELICOPTER COCKPITS

Howard P. Harper
Senior Human Factors Engineer
Sikorsky Aircraft Division, United Technologies Corporation
Stratford, Connecticut

Abstract

This paper describes the status of voice output and voice recognition technology in relation to helicopter cockpit applications. The maturing of this technology provides many opportunities for new approaches to crew workload reduction. The paper covers the helicopter operating environment, potential application areas and the impact on advanced cockpit design.

Introduction

Utilizing increasingly more sophisticated and complex on-board systems, helicopter crews will be required during the conduct of missions to perform multiple tasks which include monitoring aircraft systems, monitoring and initiating communications, navigation, target detection, air-to-air attack/coordination, active/passive defense against radar, laser and infrared air- and ground-based detectors and designators, obstacle detection and avoidance, and monitoring mission-specific subsystems.

In many cases the crews will be required to perform such tasks in an all-weather, night, nap-of-the-earth environment that demands out-of-the-cockpit visual attention and hands-on-stick control readiness. As a result of this increased task loading crew work load is approaching its maximum limit. In cockpit concepts where a single man crew is envisioned, this limit clearly will be exceeded unless a new technological approach is found. Computer voice interaction is one such approach. This paper reviews that technology and considers how it might be applied to solving some of the workload problems.

After reviewing the progress in this area it is apparent that now is an opportune time to seriously investigate cockpit applications. There are two aspects to voice technology: voice output and voice recognition. Both are currently being applied in aviation and elsewhere. In the voice output area applications range from toys and home appliances to sophisticated text-to-speech processors. The uses of voice recognition systems are not yet as wide spread but many applications are currently in full operational use. For example, they are in use for assembly line quality control and in post office mail sorting. Development of a voice input typewriter is the subject of major research efforts at several companies.

Voice Output

Techniques for producing voice output range from electro-mechanical recorders to digital sampling and storage to more sophisticated digital storage techniques such as linear predictive coding (LPC). Each of these methods has advantages and disadvantages which will be reviewed in the following paragraphs.

The best example of the use of a recorder based voice output system is the ASH-19 voice warning system used in a number of military aircraft including the CH-54 Flying Crane. Feedback from operational units has indicated that the system has functioned well over its twenty year life. It does, however, suffer from some of the reliability problems which one would expect from a complex electro-mechanical system designed in the 1950's. One drawback in an electro-mechanical system is the variation in access time to words due to positioning the playback head to the location of the next desired word.

The second technique for voice output is the use of digitally sampled and recorded voice signals. The method is simply analog to digital conversion of the speech signal and usually involves storage in read only memory (ROM). The resulting voice quality can be excellent, but depends largely on the sampling rate and encoding precision. A minimum of about 15,000 bits of storage is typically required per second of speech. The access time to words is extremely fast and as a result messages made up of strings of individually recorded words can be put together in a satisfactory manner.

Linear predictive coding was developed primarily to reduce the data storage requirements for voice output systems. This is the technique used by Texas Instruments in their "Speak 'N Spell" teaching system and in a series of chips designed to be incorporated into a variety of other applications. This technique allows storage of one second of speech with about 3,000 bits of digital memory. The result of the data compression is some loss in intelligibility when compared with a digitally sampled system. Standard vocabularies are available but special vocabularies must be processed by the manufacturer.

There are other systems available which are even more economical in terms of data storage

requirements. These may truly be called speech synthesizers because there is no recording and playback of a human voice. The speech is built up of phonemes which are the basic elements of speech sound. Using around 40 of these basic sounds along with the capability to vary pitch, intensity and timing, a synthesizer can produce understandable speech. The resulting speech has a robotic quality, but storage requirements are only about 80 bits per second of speech. While initially not as intelligible as the speech produced by other systems, it improves greatly with training and continued exposure.

The technology does exist now for the use of voice output in the cockpit environment. Considering this it would now be difficult to justify continuing the use of tone combinations as the primary auditory warning system. Defining a system will require a choice among the voice output technologies described above. Reliability considerations will probably rule out the electro-mechanical recorder. If voice quality is the primary criterion the pure digital sample and store system will likely be judged best. If, however, a large vocabulary is required one of the data compression techniques may be necessary. If a virtually unlimited vocabulary is required, as might be the case if the system were called on to output the emergency procedures now found in the flight manual, then a phoneme based system is the only practical choice.

Voice Recognition

Computer recognition of a speech input is a much more challenging problem than the production of a voice output. A variety of techniques have been used. The specific method depends on a number of variables: the size of the vocabulary, the necessary level of recognition accuracy, the number of users, the need for isolated word or continuous recognition, the time available for training of the system, and the environment in which the recognizer must be operated.

A typical isolated word recognizer works by having the user say all of the words in the vocabulary one or more times to train the system. During this process the voice signal is analyzed by a bank of filters which measure the amount of energy in a number of frequency bands. Each word is broken down into a number of equal temporal parts and the filter bank outputs for each are stored. This creates a template against which incoming words are tested. The computer finds the best match for the incoming word and carries out the appropriate action assigned to that word. There are at least ten recognizers on the commercial market today. Each claims 99% plus recognition accuracy and it probably is true that under some specific set of conditions that claim can be met. It is unlikely, however, that any of them will approach that accuracy in a military helicopter cockpit.

Current Research

Voice technology has generated a great deal of interest both commercially and in the government. Many companies are carrying out research and development activities directed toward military applications of both voice input and output technology. All branches of the military as well as NASA and the FAA have research programs in this area. There have been several conferences dealing with coordination of this work, the most recent sponsored by the Naval Air Development Center in Warminster, Pennsylvania.

The Navy has, perhaps, the longest history of military applications of this technology. They have demonstrated its usefulness in performing cockpit switching functions and in the more complex man-machine interactions of an airborne anti-submarine warfare system. NADC currently has a study under way to understand and define the problems of the Navy aircraft cockpit operating environment. This includes the effects of jet aircraft cockpit noise and the effects of G loading on the physiology of speech. This study relates primarily to the fixed-wing environment.

The Air Force is currently sponsoring a study directed toward flying a prototype voice interactive system in the F-16. This program is being conducted jointly by Lear Siegler Inc. and General Dynamics and is expected to fly this year. In the development program, progress has been made toward accommodating the unit to the jet aircraft cockpit environment. This has included dealing with problems such as the effects of the oxygen mask on speech recognition.

Helicopter Research

There are many differences in the mission and the operating environment of helicopters and fixed-wing aircraft that will have an important effect on the usefulness of voice interactive technology in the cockpit. First the missions are markedly different. The helicopter night, nap-of-the-earth, all-weather scenario imposes long duration, high workload conditions on the crew. Attention must be fixed outside the cockpit and for long periods hands cannot be taken off of primary flight controls. These conditions are often sustained for the major portion of the mission. On the other hand, fixed-wing aircraft have periods during the mission where workload is very high but these are generally of a much shorter duration.

Another factor differentiating the helicopter from the fixed-wing aircraft is the crew station environment. There are primarily two characteristics which contribute to this difference. The first is cockpit noise. Figure 1 shows typical spectra for the two aircraft types. This clearly shows the difference in frequency content. Much more energy occurs in the speech frequencies in the helicopter. The second aspect is the modulation of the voice due to cockpit vibration. This effect is shown in Figure 2 by noting the

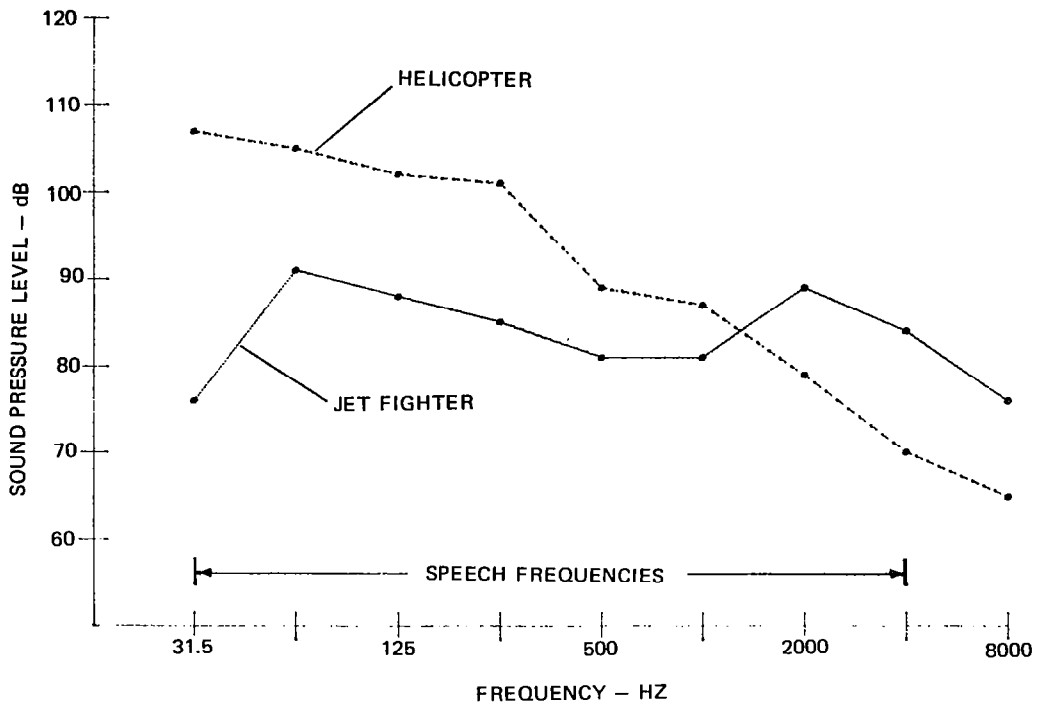


Figure 1. Cockpit Noise Spectral Differences for a Typical Military Helicopter, and a Jet Fighter

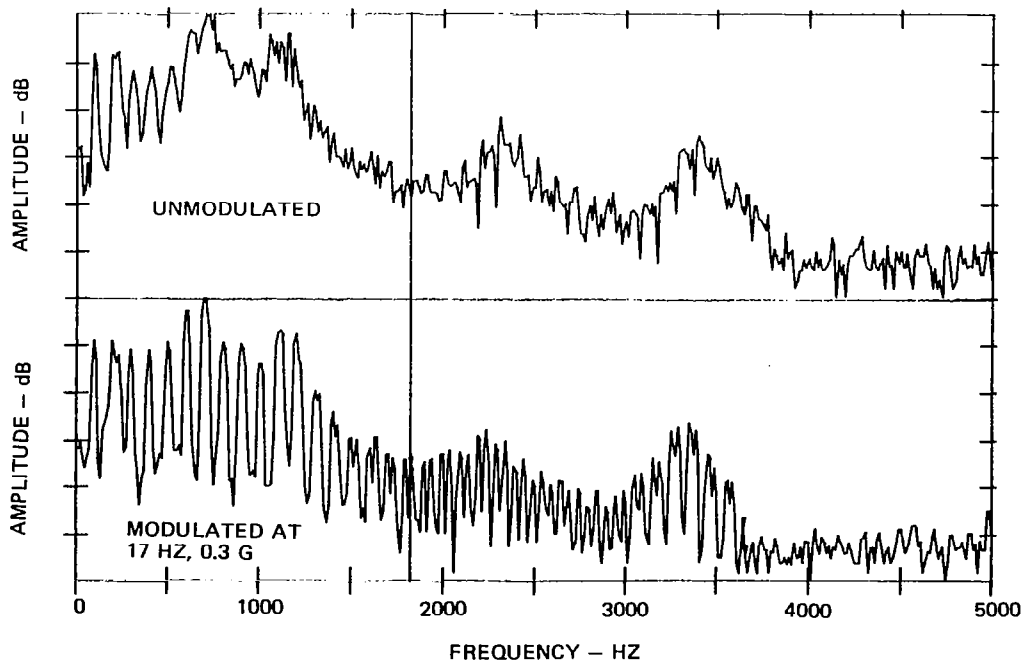


Figure 2. Modulation of a Speech Sound Due to Whole Body Vibration

difference between the voice spectra under conditions of vibration and no vibration. As might be expected, a system trained under one set of noise and vibration conditions and asked to recognize under other conditions may perform unreliably.

Helicopter cockpit-related voice technology research is currently going on at NASA Ames Research Center and at the U.S. Army Avionics Research Facility at Ft. Monmouth, N.J. The Ames facility has a long record of voice related research work. References 1 and 2 are Ames-sponsored studies relating to the cockpit use of synthetic voice warning concepts. More recent work has addressed problems of the voice recognizer in the helicopter's noise and vibration environment. The results are extremely encouraging. Even under the most adverse conditions, voice data entry compares favorably with keyboard entry. Accuracy differences never exceeded 2%. These results clearly establish the feasibility of using voice recognition in rotary-wing aircraft. In a second study currently underway, a commercial voice recognizer is being used to control an aircraft performance computer. This study is beginning to contribute information on the problems of using this equipment to perform a real function in a present day helicopter. The Army at Ft. Monmouth has taken the lead in military helicopter voice-related research. At present they are studying the noise environment of their inventory of helicopters to define the effects on the performance of currently available recognizers. Their plans call for implementation of a voice interactive system to become part of the advanced digital avionics system to be flown on a UH-60A.

Pragmatically we have to recognize that a military helicopter is far from the ideal location for a voice recognizer but, because of the work done at Ames and elsewhere, we can be reasonably certain that the problems can be solved. Therefore it should become our purpose to let the manufacturers of this equipment know that we are interested, that we can see many potential applications, and that there is a market in the helicopter industry. Furthermore we should define the operating environment so that they can do the necessary development to make equipment that will function adequately in our cockpits. Alternatively they may tell us what has to be done to our environment to make the equipment work. We will then have to address the problem of whether the value of a voice interactive system warrants the cost of an improved cockpit environment. This will provide parallel pathways for the solution of the operating environment problem and development of applications which make maximum use of the technology to reduce cockpit work.

Helicopter Applications

The following are some of the thoughts which must go into the preliminary design effort

to specify the requirements for a voice interactive system for a rotary wing aircraft. This process is needed to determine whether the time and expense of doing a complete and detailed systems and human engineering analysis is warranted.

First it is necessary to list the assumptions on which the system design will be based: 1) the availability of a speech recognizer with 100 word vocabulary with the capability of training by two users and having a demonstrated accuracy of 95 to 99.9 percent under all flight conditions; 2) a voice output device with a demonstrated intelligibility at least as good as current inter-communications systems.

Ideally this preliminary design effort would take place after the completion of a detailed analytical study of all the man-machine interactions. The results would allow evaluation of the workload reduction quantitatively and allow the designer to investigate the effects of design variables on the performance and usability of the system. The time to do such an analysis is before starting a design effort for a specific application. In the heat of a design effort the system designers cannot wait for the results of such an effort.

One of the design concepts planned is the use of the "intelligent copilot" model. All candidate voice interactive functions are evaluated in terms of whether they are consistent with the behavior of a hypothetical copilot who knows when to talk, when to listen and who prioritizes information in a logical way that is appropriate to the mission phase. A second design concept is that the system will provide feedback on all inputs and will require secondary verification of the more critical items. If, for example, the pilot were to say "Jettison Tank" the system might respond visually or orally: "Tank Jettison Requested" and the pilot would be required to confirm the request by giving an action command. Thirdly, all voice inputs are backed up with a manual entry mode which would be considered a secondary operational mode and, therefore, might require a deeper level of paging. The fourth concept is the use of a switch on the pilot and copilot cyclic grip which he will press to indicate that he is talking to the recognizer. Lastly the training of the recognizer will not be done on the aircraft, it will have been done earlier and stored on a cassette or in a ROM cartridge which can be plugged into the aircraft for a rapid data transfer.

The voice output must be unusual enough to be easily distinguished from other crewmen or air traffic controllers. This is not meant to imply that a robotic voice is required, however the voice must stand out clearly from the routine voice communication traffic. The major difficulty with robotic quality voice is that people have trouble taking it seriously and this effects its acceptability to pilots.

Next we will look at each of the various systems on the aircraft and try to understand where voice input and output technology might fit into operation of that system.

Communication

In the area of communications we will include radios for air-to-ground, air-to-air, and data links; and systems for communication within the aircraft. The functions which must be performed with this equipment include tuning, selection of the system, keying, and volume/squelch control. Tuning is a function which is particularly adaptable to a voice recognition system. The pilot might say "Tune VHF 122.7" or "Tune VHF Channel 5". Selection also fits in well with a recognizer system. The pilot would say "Select UHF" and subsequent transmissions would be made on the UHF radio. The use of voice to control volume, squelch, or keying does not seem to be practical because the voice command would interfere with the material being sent. On the voice output side, it seems possible that voice synthesis may be used to reconstruct messages encoded digitally and sent to the aircraft from the ground via a data link.

Navigation

Control and operation of navigation equipment offer opportunities where both voice input and voice output would be very effective in workload reduction. The systems which might be controlled are the doppler/inertial navigation system, Tacan, VOR/DME and ADF. The functions of this equipment are to provide: current position, steering information in X, Y, and Z coordinates, the map situation in terms of the relationship of current position to other geographical information, system updates and acceptance of flight planning inputs such as way point locations. These functions for the most part, are adaptable to voice interactive techniques. For example, current position might be called up with the voice input "Position". The system might respond in map coordinates or in terms of bearing and distance to a known point. Steering information could be requested and provided verbally. For example the pilot might ask for "Directions Waypoint 3" and the system would respond "325 Degrees, 2 Miles". Map situational information could be of the following types: request for nearest fuel or request for height and location of highest terrain in the area. Navigational system updates could easily be accomplished verbally; the pilot saying "Update Waypoint 3...Mark" when directly over the point. In addition the flight could be planned using a verbally prompted waypoint entry routine.

Flight Controls

The primary flight control system would not be directly interfaced with the voice recognizer, but system faults would trigger appropriate verbal messages. In the automatic flight control

system (AFCS) there are a number of functions which can be considered for integration with a voice interactive system. These include system turn-on, function selection, monitoring of performance, response to problems, and system shut-down. The AFCS initiate and shut-down functions are best reserved for manual action since they generally occur before and after the crew workload is at its highest. The selection of AFCS functions is a good candidate for voice actuation. Here such functions as airspeed hold, altitude hold, heading hold, or approach to hover might be selected through inputs to the voice recognizer. This is one case where a very positive feedback system would be required. A secondary command would be required prior to the initiation of any of these functions. The pilot would say "Hold Heading" and the system would respond "Heading Hold Requested". The pilot, after seeing that the system understood his input, would give an action command such as "Do It". Had the feedback been incorrect the pilot would cancel the input and try again verbally or, at his option, engage it manually. Voice output could be used effectively to provide the pilot with information on the status of the system.

Subsystems

The engine, fuel, APU, hydraulics, electrical, anti-ice and transmission subsystems might make use of voice. The possible crew functions would include system start, condition monitoring, system control, malfunction response and system shutdown. A specific engine parameter which is a very possible candidate for voice monitoring is power available. Information about power margin has a high priority at times when the pilot's attention is outside the aircraft and both hands are on the controls. The pilot might say "Power" and the system would respond with a voice message "10% Torque Remaining". Contingency power selection is a mode which allows pulling additional power from one engine when the other experiences a power loss. This selection must be set up quickly at a time when the pilot would be very reluctant to remove either hand from the controls. In the fuel system there are a number of possibilities. Voice requests could be made for fuel status with the system responding in pounds of fuel remaining or in terms of flight time remaining at the current flight condition. In addition to the low fuel warning normally provided, a programmable voice system could be used to provide a warning at any fuel state or time remaining selected by the pilot. The APU could be started and shut down by voice command but since this is generally a ground function where workload is not critical it would not be worth implementing in the voice system. Aircraft lighting is an area where a recognizer could be particularly effective. Lighting controls are numerous and frequently accessed. The voice system could select, actuate and control both interior and exterior lighting systems. In addition the voice recognizer could be used to select various sub-

system status monitor modes such as engine instruments, electrical or hydraulic parameters, or emergency procedures as suggested in Reference 3. Voice interaction with the remaining subsystems would be limited to voice messages related to malfunctions.

Caution, Warning, Advisory

The information provided to the crew by the caution, warning and advisory systems is potentially convertible to a voice output system. Those messages which are currently supplemented with an alerting tone pattern should be replaced with a voice message. With voice technology available pilots should not have to identify a failure by the pattern of tones in the alerting signal. It seems apparent that voice might become the primary alerting system for all of the warning messages and for the more critical of the caution messages. This would allow replacement of the current matrix of dedicated caution lights with a three or four line prioritized display. This type of alerting system will require some new thought because of the single dimensional quality of the auditory channel. Two messages cannot be presented simultaneously; all inputs are sequential rather than parallel. All possible messages must have a priority value which determines the order of their presentation. To complicate matters further these priorities may have to change with mission and phase within the mission.

Two recent studies (References 4 and 5) have presented conflicting data on the value of using voice warning to supplement the visual alerting system. Reference 4 found no important difference in the time required to respond various combinations of voice, tones and visual signals in a jet transport simulator. The author explains that this is because the pilots always checked the voice message against the visual caution panel before responding. The study reported in Reference 5 investigated the pilot reaction times from the presentation of a voice or light warning while flying nap-of-the-earth in a helicopter. In this case there was a dramatic improvement in response time with the voice system. It was found in this study that the pilots were willing to respond without confirming the malfunction on the caution panel because it took approximately 3 seconds to stabilize the flight path of the helicopter sufficiently to look inside. This is further indication that the helicopter and fixed-wing aircraft may require significantly different approaches to integration of cockpit voice technology.

There are several possible uses for voice input to the alerting system. One would be to acknowledge messages instead of pressing the master caution capsule to indicate recognition of the message. Another function might be to change caution priorities. If, for example, a particular system was operating marginally the pilot might want to raise its caution priority to the top of the list.

Cockpit Impact

Table I summarizes the possible voice system applications discussed in the last paragraphs. This is an exercise to identify what could be done. It is important to emphasize that the next logical step would be a thorough analysis of the functions required by the mission to determine a reasonable design solution.

The single place cockpit is the application where the need for an "intelligent copilot" is greatest. The recognizer/synthesizer will be required to take over many of the functions normally assigned to the second cockpit crewman. A single place helicopter cockpit which includes a voice interactive system is shown in Figure 3. The physical impact of the voice system is not dramatic. The only special control is the switch on the cyclic to key the recognizer. In addition, the ROM cartridge with the pilot's voice characteristics is inserted in a slot. The remainder of the displays and controls will only differ slightly from a non-voice cockpit since manual and visual backups will probably be provided for the voice functions.

The major improvements will be in the pilot's ability to keep his hands on the controls during critical flight phases, and in his capability for being fully informed on aircraft system status without bringing his eyes inside the cockpit. The concept that the recognition system will respond to simple commands will eliminate the component workload associated with finding and actuating a manual control. The use of voice actuation facilitates the use of multi-function manual controls and thus reduces cockpit space requirements to some extent.

It should be further emphasized that voice cannot be successfully introduced to cockpits on a piecemeal basis. We are beginning to see various individual systems such as ground proximity warning systems and altimeters with voice output capability. This is manageable now, but further proliferation of voice systems could become chaotic. The full benefits will only be achieved by an integrated approach.

The design of a voice interactive cockpit system requires an appreciation of the single channel nature of the auditory system. With visual displays the designer can put up a great deal of information at one time in the hope that the pilot can pick out what he needs for a particular task. With a voice system, sequencing and prioritizing of inputs and outputs is necessary since only one thing can be going on at any time.

Conclusions

The following conclusions can be drawn from this investigation of helicopter cockpit voice interactive technology:

- 1) Voice output technology is available for use now.
- 2) Research results look very favorable for the development of an accurate, reliable voice recognition system for helicopters.
- 3) There are many possible voice interaction applications which will result in workload reduction.
- 4) A thorough systems and function analysis is required to maximize benefits and to be sure that the system is acceptable to crewmen.

References

1. Simpson, C. A. and Williams, D. H., "Human Factors Research Problems in Electronic Voice Warning System Design", NASA TM X 62,464, 1975.
2. Simpson, C. A., "Synthesized Approach Callouts for Air Transport Operations", NASA Contractor Report 3300, 1980.
3. McGee, J. and Harper, H. P., "Advanced Subsystem Status Monitor, Sikorsky Aircraft", USAAVRADCOM Technical Report 80-D-5, U.S. Army Applied Technology Laboratory, Ft. Eustis, Va., 1980.
4. Wheale, J., "The Speed of Response to Synthesized Voice Messages", AGARD-CP-311, 1981.
5. Reineck, M., "Voice Warning Systems: Some Experimental Evidence Concerning Applications", AGARD-CP-311, 1981.

TABLE I. POSSIBLE COCKPIT VOICE APPLICATIONS

SYSTEM	INPUT	OUTPUT
COMMUNICATION	TUNING RADIO SELECTION	DIGITAL MESSAGE RECONSTRUCTION
NAVIGATION	POSITION REQUEST STEERING REQUEST MAP INFORMATION REQUEST POSITION UPDATE	POSITION REPORT STEERING INFORMATION MAP INFORMATION
FLIGHT CONTROLS	FUNCTION SELECTION ACTION COMMAND	FEEDBACK
SUBSYSTEMS	POWER INFORMATION REQRS. SELECT CONTINGENCY POWER REQUEST FUEL STATUS LIGHTING CONTROL SELECT DISPLAY MODE	POWER INFORMATION FUEL STATUS
CAUTION, WARNING, ADVISORY	PRIORITY SELECTION	PRIMARY ALERTING SYSTEM

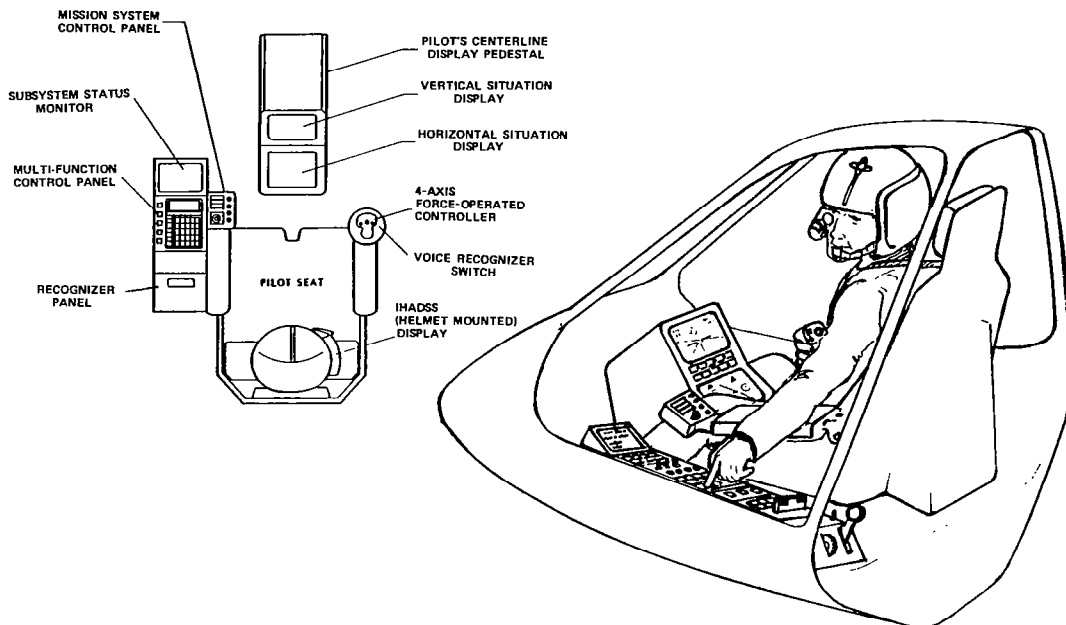


Figure 3. Single-Place Helicopter Cockpit Incorporating a Voice Interactive System