Nasa-CR-172,590

NASA Contractor Report 172590

Study To Determine Potential Flight Applications and Human Factors Design Guidelines for Voice Recognition and Synthesis Systems

FINAL REPORT

Russell W. White and Donald L. Parks

Boeing Commercial Airplane Company
P.O. Box 3707
Seattle, Washington 98124

Contract NAS1-17367
July 1985

NASA
National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

Study To Determine Potential Flight Applications and Human Factors Design Guidelines for Voice Recognition and Synthesis Systems

Russell W. White
Donald L. Parks

Boeing Commercial Airplane Company
Seattle, Washington

Submitted in Response to
NASA Contract NAS1-17367

July 1985

Foreword

This final technical report was prepared by the Boeing Commercial Airplane Company, Renton, Washington, under NASA contract NAS1-17367. It covers work performed between September 1983 and August 1984. The program was sponsored by the National Aeronautics and Space Administration, Langley Research Center (NASA-LRC). George Stienmetz was the NASA-LRC technical manager.

The authors want to express appreciation to the following individuals for discussing with us their voice applications and observations:

James McDowell, Erik Werkowitz, and David Williamson
Flight Dynamics Laboratory
Wright-Patterson Air Force Base
Dayton, Ohio

Dr. James Mosko
Rome Air Development Center
Rome, New York

Lockwood Reed
Army Avionics Research and Development Activity
Ft. Monmouth, New Jersey

Dr. Christopher Skriver and Dr. Norm Warner
Naval Air Development Center
Warminster, Pennsylvania

Dr. Victor Zue
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts

Helpful data inputs were provided by James Morris and John Wester of Texas Instruments, Speech Applications Equipment Group, and Louis Toth and George Vensko of ITT, Defense Communications Division. These individuals provided useful demonstrations, performance information, and budgetary quotes on their respective systems. We would also like to acknowledge the time taken by the 18 individuals who responded to our letter survey (sec. 2.1.2.2).

Finally, the experience and guidance of Wayne Smith, program manager, was of great value, as were the contributions of the Boeing Flight Deck Research Group members George Boucek, Donald Hanson, and Thomas Pfaff.

TABLE OF CONTENTS

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1.0 Introduction

This study was performed by the Boeing Commercial Airplane Company Flight Deck Research Group for the NASA Langley Research Center. Study objectives were to define potential commercial cockpit flight applications and human factors design guidelines for voice recognition and synthesis systems. Specific objectives were to survey existing and forecast near-term state-of the-art (SOA) voice technology, to define and appraise practicability of candidate applications for commercial flight operations, and to identify suitability for operations. A hierarchy of benefits, applications, and tradeoffs was developed. General specifications for simulator and aircraft quality voice systems have been provided. Finally, the results of this study were compared to a recently completed NASA study (NAS1-16199).

To achieve the stated objectives the study was broken down into five tasks:

Task 1: Review voice recognition and synthesis systems technology
Task 2: Appraise use of voice in control and information transfer
Task 3: Determine suitability in aircraft operational environment
Task 4: Identify and recommend applications by benefits hierarchy
Task 5: Appraise a given control-display study for candidate voice system applications

Task 1 was to define the state of the art in voice technology and forecast its progress over the next five years. This effort would thus produce the baseline tradeoff data for identifying practical applications of voice recognition and synthesis systems for aircraft operations.

Task 2 was to review aircraft system management requirements for possible uses of voice systems. This included identification of potential uses in existing systems as well as possible uses in evolving systems.

Task 3 was to synthesize environmental and operational considerations for flight deck applications. This effort was to identify both benefits and constraints of candidate voice systems applications identified in tasks 1 and 2 for practicality and desirability of use for each potential task function.

Task 4 was to develop a hierarchical benefit scheme for the applications under consideration, which could be used to contrast voice systems concepts versus traditional hardware systems concepts. It would also develop general-purpose human factors guidelines for voice system uses in the flight deck and general specifications for three levels of candidate voice system capability.

Task 5 was to appraise another NASA control display study for possible insertion of voice system applications.

2.0 Technical Report

2.1 Task 1: Review Voice Recognition and Synthesis Systems Technology

The objective in task 1 was to define the state of the art for voice recognition and synthesis technology, project where it will go in the next five years, and establish a baseline capability for aircraft and simulator applications of voice input-output technology. A literature survey helped to determine the capabilities and performance available in current voice systems. The literature survey information was supplemented significantly by surveys of voice system manufacturers, expert and general users, and an academic research center. A summary of resulting indications and conclusions follows.

2.1.1 Voice Recognition-Synthesis Overview

Voice recognition is still awaiting large-scale commercial applications. Most systems are being purchased for evaluation purposes. A few applications in industry and government have been reported, including baggage handling, quality control, and record keeping functions, especially for functions where the hands and eyes are busy. Capabilities range from recognizing single words or phrases separated by definite pauses to interpreting entire sentences spoken in a natural manner.

All the systems, with exception of a few with very limited vocabularies, have the disadvantage of requiring the user to "train" on the system by speaking each word to be recognized one to ten times. This training requirement makes the system speaker-dependent. Another disadvantage is that the number of words that can be recognized at one time and acknowledged in real time ranges only from 30 to 100.

Determining (or predicting) the accuracy a voice recognition system will actually demonstrate in a working environment is difficult but very important. Also important is the type of errors possible: rejection of a vocabulary words; misidentifying a vocabulary word and substituting another; or misidentifying a sound or nonvocabulary word for a vocabulary word. A user must determine what are acceptable accuracy rates and types of errors for a particular application and environment. Potential systems will probably have to be tested by the user for compliance because this type of information is not available from the manufacturer.

An Air Force-Navy-NASA research group has been working with domestic and foreign voice systems manufacturers to develop a flight quality voice recognition and synthesis system (reference sec. 2.1.2.3). The systems had to operate in the demanding environment of an F-16 fighter aircraft cockpit. Two systems flew in the Advanced Fighter Technology Integrator (AFTI) F-16 test bed and two others showed technical feasibility to do so. The four systems had limited vocabularies and only operated in isolated word recognition mode, but they all operated with reasonable success in the high ambient noise environment (115 dB) of the F-16.

Several of the voice systems manufacturers have already begun improving their systems to handle limited connected word recognition. The connected word recognition systems should be ready for military and commercial use in one to three years. Vocabulary active at any one instant for a given use will be about 40 words, but total potential vocabulary can be several hundred words. Flight qualified speaker independent voice recognition systems will not be available in the next five years.

Voice synthesis (or response) systems are used in a wide range of applications from children's toys to automobiles. Some aircraft quality systems are available and are in use in commercial and military airplanes.

Voice synthesis systems fall into two categories: those that play back prerecorded (digitized and compressed) words and phrases, and those that create words by combining various phonemes.

The digitized and compressed synthesis systems provide the highest quality voice reproduction. The few flight quality systems are of this type, as are the bulk of commercially available voice synthesis systems. The drawbacks of this technique include committed vocabulary and having to prerecord each word.

Phoneme-type voice synthesis systems are commercially available now, but flight quality systems have not yet surfaced. Some phoneme systems are barely understandable, but a few offer good quality and various voice types; child or adult, male or female. Predefined vocabularies of greater than 7,000 words are available. It is technically feasible to ruggedize some of the better quality phoneme systems now; development depends on availability of a market.

Terminology for voice recognition and synthesis systems characteristics is provided here to acquaint the reader with some of the peculiarities of voice systems related to the range of capabilities.

- **Training:** Most present recognition systems require "training" for the accumulation of a statistical record of voice characteristics. Each user must train a particular system before it will recognize that user saying a number of predefined phrases or words.

- **Limited Training Recognition:** Training involves reciting a few predefined word strings to the system. The number of words recited would be a small subset of the total number of words the system could recognize.

- **Isolated Recognition:** A user must provide a definite separation between recognizable (trained) words or phrases, e.g., 200 milliseconds. Often the user must wait for the system to acknowledge that a phrase/word has been recognized or rejected before entering another phrase/word.

- **Connected Recognition:** A user may speak a number of recognizable (trained) words without pausing between them. In connected word systems, the words must be said carefully and not spoken quickly (National Bureau of Standards, ref. 14). The number of words that may be spoken and recognized varies from a few to unlimited, depending on the system. Interjecting nontrained words may cause the system to signal an error or hang up.

- **Continuous Recognition:** A user may speak a number of recognizable (trained) words in a fluent manner and at a natural speed. The number of words that can be recognized at one time is dependent on the system. Interspersed unrecognizable words may cause errors.

- **Word Spotting Recognition:** Trained words that are interspersed in word strings with untrained words can be recognized selectively, and the untrained words will be ignored. If the system is capable of connected or continuous recognition, two or more trained words may be strung together so that the system selects words in context from the string of words an operator might say.

- **Speaker-Dependent Recognition:** Each user must train a particular system before it will recognize the specific user saying specific phrases or words. Training involves speaking each phrase or word that the system is to recognize one or more times.

- **Speaker-Independent Recognition:** The system will recognize a large percentage of a population saying certain phrases/words without significant loss in accuracy. The population will most likely be linguistically common.

- **Variable Syntax:** The ability to correctly respond to key words regardless of the order of presentation.

- **Syntaxing a Recognition Vocabulary:** The ability to direct a system to recognize a specific subset of the total vocabulary it has resident in RAM or ROM memory.

- **Digitized Synthesis/Playback:** Phrases, words, or word segments are prerecorded into a memory device in digital format. Often compression techniques are used to conserve memory. Voice synthesis/playback involves pulling a words record from memory and sending it through a digital-to-analog converter, audio amplifier, and a loudspeaker or headphone.

- **Phoneme Synthesis:** Phonemes, the smallest distinguishable unit of speech utterances, are stored in a systems memory. Words are created by combining two or more phonemes according to predefined rules (resident in a system's memory) or commands from a host system.

- **Text-to-Speech Synthesis:** ASCII code received from a host system, a terminal, or an optical character reader is converted into a word or word string. The text-to-speech system has resident firmware that is used to relate the ASCII code to phonemes that in turn form words.

- **Linear Predictive Coding (LPC):** A common method of extracting pertinent features of speech for voice recognition, synthesis, or voice encoding/recording. The LPC method represents a "speech signal in terms of the parameters of a filter whose spectrum best fits that of the input speech signal." (Doddington and Schalk, ref. 10.)

2.1.2 Survey of the State of the Art in Voice Systems — Existing and Near-Term Capabilities and Tradeoffs

Major changes in voice technology have occurred in the last five years, with the most marked changes occurring in the last two years. Part of this is indicated by the large number of current manufacturers (Appendix A). Also, recent reports show a shift from basic research to a significant amount of applications research, a trend that started around 1980.

Surveys of over 250 manufacturer and user organizations have produced a list of 62 systems that are available for application, with information on their performance characteristics. Systems vary widely in cost, capability, apparent ability to operate in "unfriendly" environments, and extent of associated experience. Few are recognized to be sufficiently ruggedized for aircraft use. Five systems were built to military criteria and are being tested to demonstrate adequate performance in the very severe environments (115 dB and 5 g_n) of the Advanced Fighter Technology Integrator (AFTI). Others may have been designed to a standard of rough use sufficient for less severe nonmilitary airplane environments, but data made available does not permit a judgment. Certainly, those designed to military criteria could be expected to have more than adequate capability for present purposes.

This 1980 distinction in technology analysis is also evidenced by a shift in emphasis in available literature, with a rather large proportion of applications-oriented articles being published in the brief time since 1980. Overall, of 1,956 articles generally related to the field, only 357 were relevant to current purposes, of which 114 were published since 1980. These latter references are included in Appendix B; for convenience, they are in two sections, post-1980 and through 1980.

Additionally, selected but key technology centers were visited for face-to face discussions and critiques of capabilities, limitations, and projected resolutions with technical experts in the field. There was concurrence among these experts that many key voice systems problems have been resolved for present objectives, that is, applications in the relatively benign environment of commercial aircraft.

Finally, major voice systems state-of-the-art appraisals are being conducted in the AFTI program. Status of that program was provided to support the present effort.

2.1.2.1 Literature/State-of-the-Art Survey Results

Numerous good literature reviews already existed (refs. 8, 17, and 18), and a major update was provided in the 1983 American Voice Input/Output Society (AVIOS) proceedings (ref. 1). Accordingly, review of the literature for present purposes was constrained to identification of relevant information regarding use of voice in control display applications for aircraft operations. The choice herein was to summarize user-oriented information rather than present a theoretical review and critique of the literature. The intent was to identify candidate applications and both benefits and constraints of importance for voice system operations rather than process a traditional literature review.

- Overview of Voice Systems Applications

Voice input can be used for switching; however, it doesn't appear to offer much advantage over traditional switching operations. An access word, an action word, and feedback would be required to complete and confirm a given actuation. The process does not offer apparent advantages over traditional switching, which inherently provides access and operation cues from location, functional grouping, switch position, and force-deflection actuation cues. However, there may be benefits if the task requires that visual attention **must** be addressed elsewhere, or when **continuous** speech recognition is possible.

Given some degree of interaction with the displays (particularly electronic display), the scope of potential uses for voice recognition can change significantly. Checklist operations, control of multifunction displays, and multifunction switching could all be performed by voice input. More sophisticated tasks, such as systems and data management, are possible — to recall, update, and enter new information. Automated communication to transfer messages (for instance by data link) becomes possible using voice to format the message, display a menu to coach and/or copy, to confirm as required. Voice programming to reprogram the flight plan or flight management systems is also possible.

Voice synthesis could be used to provide feedback (in a confirming role for inputs) or in a more dynamic interactive role, much like a dynamic visual display. Candidate applications for voice synthesis include alerting, auditory data display, interactive functions (such as checklists, procedural recall, or data link message transfer), and training. However, there are potential conflicts in use, demonstrating that applications should be made with caution and synthesis should be used sparingly as concluded in an extensive FAA-sponsored study to develop caution-warning guidelines (ref. 6). Conflict sources include alerting signals (used to demand and focus attention, prioritize urgency, and guide procedures), interphone communications, and radio communications. A concern is the degree of use of the auditory channel; the single thread nature of the channel requires that new applications concepts minimize the potential for message interference or selective attention that filters out information. Additionally, while auditory information is demanding it is not as readily selected or reappraised as visual displays. There is also concern regarding repetition of similar signals: flight crews might acclimatize and become conditioned to ignore many auditory messages.

- Voice Systems Applications Problems

Remaining technological problems to be resolved for voice systems operations are more extensive for voice recognition systems than for voice synthesis systems. Voice synthesis is relatively straightforward, with all pertinent variables under design control. However, many variables in voice recognition are not so easily resolved — the design must accommodate wide variations in speakers, in characteristics such as speech time, inflections and enunciation, and in distortion sources such as stress and fatigue as well as interference from ambient noise. The sources of variability are factors in processing strategy and computation time, which, in turn, can influence the applicability of a given system.

- Voice Recognition Systems Constraints

Processing time and strategy are significant factors that could affect operational use and acceptance of voice recognition systems by an aircrew. Early systems were unacceptable except for demonstration; they operated in nonreal time and required an unnatural hesitation between words. Newer systems have improved the algorithms to give a much faster response, but potential use can be restricted by the strategy for the algorithm.

The most common constraint is that of accommodating acoustic similarities between words. Most systems require speaker training to obtain a pattern match between programmed words and speaker speech characteristics; current speaker-independent systems are only available for very small acoustically dissimilar vocabularies.

Next, uninterrupted continuous speech, as with the spoken language, is not yet possible. Some systems have the appearance of continuous speech because they are only processing trained phrases or picking up trained words in context out of a conversation. Others operate in nonreal time, waiting for completion of a sentence, then using verbal characteristics of the whole sentence as part of the processing algorithm. Still others use a syntaxing capability, collating permissible syntaxing with the sounds of the words and phrases to reduce emphasis on individual word recognition. Also, vocabulary and syntax restrictions are used to reduce the amount of processing by limiting the amount of recognition required at any one time.

An important constraint is recognition accuracy. Accuracy can be affected by such factors as vocabulary size, acoustic similarity of words, and the number of operations processed. Accuracy figures are a function of several factors, so accuracy discussions and figures can be confusing. One measure of accuracy that is most often quoted is based on using a sound recording tape to train the recognizer, then using the same tape in playback as a test of recognition accuracy. Unfortunately, such accuracy quotes seldom consider the normal variations in the spoken word. Another accuracy factor relates to the threshold settings for acceptance or rejection of words. Error sources can include false acceptance of an incorrect word, false rejection of a correct word and substitution of another (acoustically similar) word for the one actually spoken (gear vs. cheer, A J vs. H A). Such error sources can be modified by changing the rejection threshold and thus the trade-offs of accuracy, false acceptances, and substitution. Certainly, for most aircraft applications it would be better to reject wrongfully than to accept a wrong word. For this reason, some demonstrations provide for preentry screening and confirmation before permitting an entry or action.

Variations in operator characteristics present another constraint of use. The importance of distinct and consistent enunciation has led to categorization of speakers as "sheep" and "goats." Some people (sheep) are highly repeatable in their voice pattern, thus are naturally qualified for ready use of voice recognition systems. Others (goats) are not consistent from time to time, making voice pattern matching more difficult. Accordingly, systems based on pattern matching (the trained/acoustic pattern) have a more difficult time recognizing the words. Similarly, conditions that contribute to voice distortion (presumably stress, fatigue, hoarseness, and similar factors) may disrupt voice recognition accuracy.

The ambient noise environment can have a definite negative impact on a recognizer's performance. When a system cannot differentiate the ambient noise from the user's voice it may shut down or produce a large number of errors. Prescreening of the voice input by special hardware can control environmental interference from ambient noise or background talk. Using noise-cancelling or directional microphones is a simple method to control such interferences. Microphone distance is also important; a fixed distance is necessary to maintain voice amplitude vs. ambient interference. In application, added control can be introduced by use of a mike-keying switch, so that the user consciously opens the mike channel when he wants to talk.

More elaborate methods are also possible. For example, at Rome Air Development Center, signal enhancement techniques are being used to improve the signal-to-noise ratio. They have demonstrated a 15 dB improvement in speech signal-to-noise level, which made the difference between an unintelligible and an intelligible message. More extensive emphasis on preprocessing of the signal, before recognition is attempted, is apparently receiving more attention.

Also, training in the ambient noise environment can facilitate recognition of the spoken word; an Interstate Electronics system recognized words when trained with a 106 dB factory noise environment (ref. 11). Obviously, isolated word recognition with a definite pause between words (e.g., the Interstate system) can work in extremely loud ambient noise. The clearly defined end points make it easier for the system to distinguish the voice from the ambient background. Connected speech or continuous speech systems do not need the break, but it becomes more difficult to judge end points in a noise environment since they work with the entire word string.

The requirement for user training on speaker-dependent recognition systems is a definite constraint. Early voice recognition systems required intensive training in the environment to achieve reliable operation. Now, as little as one pass through the vocabulary can be sufficient to achieve a creditable demonstration, although more passes might be desirable for improved reliability. Other concepts are also being considered in order to simplify the user load in rapidly achieving an adequate level of training for system operation. One concept is that the user would have a personal tape or magnetic card that would have a prerecorded vocabulary to present the user's voice characteristics, encoded to significantly reduce the training cycle. Another concept would utilize a simplified paragraph of phonemes or words that could be read or typed for training of the system, which could be self-adaptive with feedback from continued use or updated with a fixed phrase. These latter features would accommodate for day-to-day and seasonal variations in the voice.

- User Constraints/User Conveniences

Newness of voice recognition systems is such that most researchers have addressed the recognition technology problem and few have addressed the user interface — user operations problems. Some major constraints still exist. However, some key characteristics to enhance this interface and enhance interactive operation are emerging.

The most commonly recognized constraints relate to the fundamental operation of the voice recognizer. Examples include the forced delay of speech input imposed on the user by systems that require discrete enunciation. Also, there is the tendency of some machines to stop or do a substitution when they do not recognize a word.

Another limitation is the rigid syntax imposed by some systems, which in turn requires a very strict structuring of word sequences. Of course, such limitations are receiving more attention as the state of the art evolves and researchers turn to concept refinements. Some of the most useful features now becoming available are in syntax provisions. Some systems have a syntax feature called word spotting — the recognizer selects the programmed action words from spoken phrases or sentences. Others have a variable syntax capability which avoids strict sequential structure for words and permits the type of interchange one might normally perform, for example, Portland ILS approach; ILS approach, Portland; or Portland approach, ILS. Variable syntax also offers the potential to directly access an area of interest rather than labor through a menu series or through a series of switching procedures. Rapid access to the ultimate indenture could facilitate numerous operating procedures.

Meaningful user interface conveniences should evolve with the interactive input-output (I-O) applications which are receiving increased attention as a means of improving man-machine interface operations. Both auditory and visual feedback modes are being demonstrated. U.S. Army research is exploring an integrated approach via recognition and voice synthesis systems, using voice synthesis in both a feedback and in a display mode. Also, Boeing has been exploring integration via recognition and visual display systems, using a visual display to confirm input and also using integrated formats for the visual displays, to improve comprehension and to restrict the user less (ref. 21). Several possibilities for increased efficiency can be envisioned, especially if combined with a line edit designator, a tracker-designated crosshair, or a touch sensitive display that can be used to instantly access the appropriate display area for information of concern. For example, the touch point (or crosshair) can access the specific part of the displayed information that is the data point of interest — which could be indicating a system fault (perhaps via color code) — then verbal change instructions can be spoken, and synthesized verbal or visual feedback can be provided regarding the system interpretation of the instructions before an execute command is given.

- Review of Applications and Research Projects

Voice system applications questions are receiving more attention now that system performance has progressed to the point where operational uses are being seriously considered. As a sign of the level of maturation, some of the uses identified in papers and discussions at the 1983 conference of AVIOS are listed in Figure 2.1.1. Admittedly, some of the applications are experimental, but they demonstrate a level of interest and of confidence in utility.

Supporting the applications has been an increase in user oriented research. A brief summary of selected user-oriented research projects that have been accomplished for voice systems will convey progress in this area. The summary is organized to give synopses by categories. In particular, it will be noted that an extensive amount of this research has been accomplished at the Naval Postgraduate School and Indiana University.

a. **Accuracy:** Acceptable accuracy rates for voice systems can make the difference between acceptance and rejection. For example, if 95% accuracy is adequate, many available systems are useable. Conversely, if 95% is inadequate, the cost in bad reputation could seriously damage voice systems progress. Rates of 99%, 95%, 90%, and 85% were explored. Results were not conclusive. (Poock, G. K. and Roland, E. F., 1982, ref. 26.)

b. **User Experience/Speaker Independence:** Both naive and practiced speakers achieved 96% accuracy in a speaker-independent mode with a Threshold Technology T600 system trained by 10 passes of 50 utterances, by four voices other than their own. Both groups attained about 96% accuracy. Nonrecognitions (substitutions) accounted for 70% of the total errors. (Poock, G. K. and Martin, B. J., 1983, ref. 29.)

Voice Recognition/Synthesis Applications: Test and Operational Uses

Sales order taking

Telephone
- Interaction
- Dialing

Reservations

Software programming
- Computer programming
    - Keyboard entry substitution
    - Output
- Data processing equipment management

Information retrieval systems

CAD/CAM applications (computer aided design/computer aided manufacturing)
- CAD graphics
    - Development
    - Mod—"put that there" operation (e.g., with light pen interaction)
    - Cartography/geodesy editing/charting
    - Source data entry automation
- CAM
    - Manufacturing systems machines/processing control
    - Inspection/quality control
        - Records entry
        - Task simplification

Text applications
- Text to speech (with optical character reader)
- Type to speech translation
- Speech to text translation

Air carriers
- Baggage/package processing

Business
- Data access terminals
- Electronic/dialing
- Products application
    - Home electronics
    - Automotive
- Inventory control
- Automatic sorting

Banking
- Automatic teller
- Credit card recognition

Medicine
- Drug manufacture
- Medical records
- Anethesia record keeping
- Handicap support devices
    - Robotics control by limb disabled
    - Input/output use by blind

Security
- Monitor—alert
- Entry access recognition/verification (<2% errors)

Figure 2.1.1. Reported Applications of Voice Recognition and Synthesis Devices From 1983 American Voice Input/Output Society (AVIOS)

Of course, for aircraft systems operation, all backup control systems could be kept concurrent electronically. For example, syntax provisions allowing voice input to menu bypass for direct access to a detailed function could bypass some of the steps a multifunction switch (changing function-changing legend) or a keyboard would normally require. However, the multifunction switch or keyboard system could be kept concurrent with the voice activated changes in order to maintain current status for the manual backup mode if required. Possible applications include radio settings, message format for data link, and system management monitoring and adjustments.

Voice input-output formatting requires more research and development for aircraft operations. Response time, error potential, and workload can all be significantly affected by the design of the voice interactive system. Simplified input requirements, meaningful feedback, and reliability of operation are examples of areas needing better research data for design provisioning. Additionally, the operations situation requires attention; for example, key guidelines resulting from caution/warning standardization studies were to minimize auditory messages in the flight deck in order to avoid disruption of warnings, air traffic control messages, and crew communication.

c. **Voice Change Over Time:** Exploration of voice recognition performance as a function of time showed no decrement after 21 weeks (2.07% errors). This study was on eight subjects using their own voice patterns for vocabulary sizes from 20 to 240 utterances. 57.37% of these errors were by two people. In an added exploration, an error rate of less than 2% resulted for trained voices when combining voice reference patterns of two people (male and female). (Poock, G. K., 1981, ref. 24.)

d. **Use of Masks:** Measuring effects of a stenographer's mask on recognition accuracy rates showed 2.4% errors compared to 0.4% without masks for speakers experienced in using masks or microphones, and in experienced speakers showed 7% errors with masks compared to 1.6% without masks. (Poock, G. K., Schwalm, N. D., and Roland, E. F., 1982, ref. 25.)

Voice recognition with a speaker's mask varies with the type of mask and microphone. Using an Interstate Electronics VRT 101, error rate increased when using a stenographer's mask, but the increase was not of practical consequence since total error remained under 2%. However, average error rates for gas mask users increased markedly from the no mask average of 7.7% errors to 12% using a dynamic microphone and 9.2% with a noise-cancelling microphone. Gas mask effects were also shown in a one subject exploration with the Threshold Technology T600. The authors suggested that user experience with microphone and mask and better placement of the microphone in the mask could produce accuracies similar to the "unmasked" condition. (Poock, G. K., Roland, E. F., and Schwalm, N. D., 1983, ref. 28.)

e. **Speaker Independence:** Using the Threshold Technology T600 voice recognition device, an experiment addressed the possibility of achieving speaker independence using speaker-dependent equipment. Recognition accuracies of 95% resulted from storing four user patterns without the present speaker; accuracy increased to 99% when the speaker's pattern was stored along with four other users. (See also c, above.) (Poock, G. K., et al, 1982, ref. 27.)

f. **Language/Accent Independence:** In a bilingual comparison using the Threshold Technology T600, voice recognition was roughly equivalent for each test language, but recognition performance was severely degraded when the two languages were combined for simulation of reversion to native language by the speaker. The difference was attributed to the resulting complex array for the two languages that was required for interchangeable voice recognition. (Neil, D. E. and Andreason, T., 1981, ref. 22.)

Flight plan filing was accomplished by 15 subjects with regional dialects over a telephone, with messages received by an utterance recognition device that had been trained by 5,617 voices from 24 cities of the U. S. to assure dialectic representation. Ten of 15 subjects were able to file flight plans that were 100% accurate. Accuracy for the other five ranged from 93.2% to 99.4%. (Skochet, E., Quick, P. and Delemarre, L., 1982, ref. 32.)

g. **Voice Changes—Stress, Sex, and Characteristic Differences:** Task-induced stress led to more monotonous speech, speech pattern irregularity, and differences in volume, fundamental frequency, shift in fundamental frequency, and articulation precision. (Heckey, L., et al, 1968, ref. 15.)

Voice entry error rates using a Threshold Technology T600 unit showed no difference between male and female or officer and enlisted personnel. (See also c, above.) A training plateau was achieved within five training passes; there were more errors between three passes and five, but no difference in error between five passes and ten. (Batchellor, M. P. and Poock, G. K., ref. 5.)

h. **Data Entry:** Precision entry requirements for a Warfare Environmental Simulator showed that voice entry performance of the task in the required highly formatted, error-free fashion was too long and resulted in too many errors on the Threshold Technology T600 unit. Typing produced fewer errors, but typing and buffered voice showed no statistical difference in time for the same ultimate accuracy. (McSorley, W. S. III, and Poock, G. K., 1981, ref. 19.)

Voice data entry for reporting of imagery intelligence for command, control, and communication operations using a Threshold Technology T600 resulted in 97.0% accuracy without rejects of words that were in doubt and 95.5% with rejects (which also included some recognized words). Comparison of buffered voice and unbuffered voice with typing showed, respectively, 58% and 41% faster entry. Voice was as accurate as typing for writing short "order of battle" reports. (Jay, G. T., and Poock, G. K., 1981, ref. 16.)

Comparing data entry for stores management and navigation preflight for a P-3C using the Threshold Technology T600 voice recognizer versus the standard keyboard normally used, voice entry was faster for stores management (multicharacter entry) and slower for the navigation preflight tableau task (character-by-character entry). However, subjects with prior voice entry experience did better on both. (Taggert, J. L., Wolfe, C. D., Neil, D., and Poock, G. K., 1981, ref. 34.)

Use of voice systems to input cartographic data showed voice to be easier, faster, and more accurate than the paper, pencil, and keypunch method being used, plus it eliminated the need for skilled typists. (Scott, P. B., 1978, ref. 31.)

Two major experiments explored applicability of voice systems in air traffic control, using a Threshold Technology T600. (1) 12 operators spoke 46,000 words of operational data entry language with only 1% error. Algorithm refinement reduced the error rate to 0.4%. (2) Five operators using voice or keyboard entered 6,000 messages from 24 basic message types; voice recognition produced 64% fewer errors than keyboard entry. For most messages, there was no difference in rate of entry, although overall rate showed a marked advantage for voice recognition since field delimiters (or "punctuation" format encoding) were required for keyboard entry. However, slower voice recognition of digits produced a keyboard advantage for digital entry of messages; messages that were mostly digits were 30 to 50% faster with a keyboard, although with less visual confirmation since the operator always looked at the keys. Finally, feedback was faster and correction less difficult with visual than with auditory feedback. (Connolly, D. W., 1979, ref. 9.)

i. **Training:** Training method and experience level of speakers were key characteristics in recognition accuracy for a Threshold Technology T600 compared to lesser but recognizable effects related to computer experience, accent, time of week, vital capacity and rate of air flow, speaker cooperativeness, and anxiety. However, it was concluded that practically anyone may be a potential candidate to operate voice equipment. (Yellen, H. W. and Poock, G. K., 1983, ref. 35.)

j. **Workload, Speed, and Stress:** Speech recognition with an isolated word system was as fast as a visual-manual entry mode when using only one- to two-word utterances. However, for a string of digital entries, speech recognition with an isolated word system resulted in lengthened transactions compared to keyboard entry. Voice system benefits were indicated for more complicated tasks. Simultaneous tracking performance was least affected when voice recognition (isolated word) was used for other input tasks, being best when interactive speech (i.e., with audible feedback) was used for entry modes, next best when speech input was used with visual feedback and least best when manual input with visual feedback was used. (Mountford, Schwartz, and Graffunder, 1983, ref. 20.)

Voice recognition performance degraded as various mental loading conditions, involving increased short-term memory demands, were imposed on operators. (Armstrong, J. W. and Poock, G. K., 1981a, ref. 3.)

Mental loading involving decision making led to degraded voice recognition performance of two types: (1) Recognition error rate increased quickly during the first five minutes, then increased at a slower rate; (2) subjects' verbal error rate also increased with load. Compared to earlier work by the same authors, any amount of mental loading seemed to degrade performance, but there was no difference in performance between the various amounts of mental loading. (Armstrong, J. W. and Poock, G. K., 1981b, ref. 4.)

Increased motor loading via a tracking task paralleled degradation in voice recognition system performance, indicating a stress effect from task load on speech enunciation. Mean error rates were 39% greater during voice tasks with concurrent tracking versus without concurrent tracking. (Armstrong, J. W., and Poock, G. K., 1980, ref. 2.)

Using voice input versus manual typing, and with minimal practice, a command and control task was done 17.5% faster, and 25.0% more was done on another task that was being performed simultaneously. Manual typing had 183.2% more entry errors than voice; however, observers were more critical of voice system errors than of typing errors. (Poock, G. K., 1980, ref. 23.)

Stress was induced by reducing message time as subjects spoke to a Threshold Technology T600 recognizer. Voice recognition rates decreased as time was reduced, although accuracy was maintained at 90% for the worst case stress. (French and Poock, 1983, ref. 13.)

2.1.2.2 Letter Survey

At beginning of this study, a developer/user survey was conducted to determine how voice systems were performing in the field. Based on publications and contacts with users, a number of contacts were identified and polled by mail. The survey was undertaken to augment performance data for information relevant to users, covering such factors as different acoustic environments, user training and experience, etc.

Over 250 survey forms were mailed to academic and industry voice system developers and users of voice recognition and synthesis systems, who were requested to indicate on the forms which voice systems they had used, briefly note each system's features and performance, and add any pertinent comments. A five-year forecast was also requested. Two-thirds of the completed forms were from voice system manufacturers and the rest from independent research laboratories. Results of the letter survey are presented in Figures 2.1.2, 2.1.3, and 2.1.4. Overall, 62 systems were identified by the various survey efforts (Appendix A).

The survey demonstrated, first, that there appears to be little standardized performance data available for voice recognition and synthesis systems. Second, as reported by the independent users, the performance and ease of use of voice recognition systems was limited, although personal discussions with some users (and papers presented at the 1983 AVIOS conference) indicated there are a number of satisfactory applications. Third, perhaps the most important item learned from the survey was some of the qualifications that should be considered for any voice data system, for example:

- The environment the voice system was expected to operate in, that is, type of mix and ambient noise level
- The system training users receive
- Experience users have with voice systems
- Size of vocabulary used and basis for selection
- The partitioning of the recognition vocabulary

Such factors have a significant impact on the performance of voice systems in terms of accuracy, type of errors experienced, and causes of errors.

VOICE RECOGNITION SYSTEMS

FEATURES

(Survey Response)	MANUFACTURER, LOCATION	DEVICE(S)	PRICE $K	MODE (ENTER ✓): SYNTHESIS OPTION	MODE: CHIP	MODE: BOARD	MODE: SYSTEM	MODE: SOFTWARE	AVAILABILITY (ENTER ✓): RUGGEDIZED	AVAILABILITY: MILITARY QUALIFIED	AVAILABILITY: MILITARY QUALITY	AVAILABILITY: COMMERCIALLY AVAILABLE	AVAILABILITY: PROTOTYPE	AVAILABILITY: DESIGN	AVAILABILITY: SUPPORT SOFTWARE	TYPE (ENTER ✓): DISCRETE WORD	TYPE: DISCRETE WORD, RAPID INPUT	TYPE: CONNECTED WORD	TYPE: CONTINUOUS SPEECH	TYPE: SELECTED WORD IN CONTEXT	TYPE: SYNTAX CAPABILITY	USER INTERFACE (ENTER ✓): SPEAKER DEPENDENT	USER INTERFACE: SPEAKER INDEPENDENT	USER INTERFACE: SPKR VERIFICATION SECURITY	NO. OF WORDS (ENTER NO.): BASIC NUMBER	NO. OF WORDS: OPTION NUMBER	NO. OF WORDS: UTTERANCE LENGTH (SECONDS)
(1)	NEC	DP-200	14				✓					✓				✓		✓			✓				50	250	
(1)	VOTAN	V5000	5				✓					✓				✓					✓				256		4
(2)	VOTAN	V5000	5				✓																				
(2)	Interstate	VRT 101					✓																				
(2)	Interstate	EVL 008			✓																						
(3)	Super Soft, Inc.	Scratch pad with Voice Drive + TECMAR voice recognition board	.995			✓		✓				✓				✓						✓					
(3)	Super Soft, Inc.	Scratch pad with Voice Drive	.495					✓				✓								✓		✓					
(4)	VOTAN	V5000A	6				✓					✓				✓					✓	✓			256		1

PERFORMANCE BASED ON EXISTING OR READILY AVAILABLE DATA

(Survey Response)	MANUFACTURER, LOCATION	DEVICE(S)	NOISE EFFECTS (ENTER NO.): ACCURACY WITH DIRECTIONAL MICROPHONE: % STD TAPE TEST	DIRECTIONAL: % QUIET ROOM (55 db)	DIRECTIONAL: % MAX Db FOR SAME	ACCURACY WITH NOISE CANCELLING MICROPHONE: % STD TAPE TEST	NOISE CANCELLING: % QUIET ROOM (55 db)	NOISE CANCELLING: % MAX Db FOR SAME	(OTHER): % STD TAPE TEST	(OTHER): % QUIET ROOM (55 db)	(OTHER): % MAX Db FOR SAME	USE/SUSCEPTIBILITY: ACCURACY IN %: QUIET AREA	ACCURACY IN %: IN USE AREA	ACCURACY IN %: TRAIN IN NOISE	ENTER RATING (1 POOR THRU 5 GOOD): EASE OF ADPT./IMPLMTN	TRAINING EASE	PERF. CONSISTENCY	AMBIENT NOISE PERF	PULSE NOISE PERF	MIKE DISTANCE	USER ADAPTATION	VOICE VAR (DAILY COLD ETC.)	LANGUAGE DEPENDENT (YES ✓ NO 0)	COMMENT (INCLUDES WHO'S CHIP, SPECIAL EQUIPMENT, SIZE/VOLUME, WEIGHT, METHOD, WHO'S COMPUTER ETC.) & FIVE YEAR FORECAST (SYSTEMS CAPABILITIES ETC.)
(1)	NEC	DP-200													3	5	4	1					0	
(1)	VOTAN	V5000													5	5	5	3					0	Standard run mode
(2)	VOTAN	V5000		>97			N/A			—		>97			5	5	5	4	3	4	5	5	0	Very good unit
(2)	Interstate	VRT 101		70			75					70			4	3	1	2	3	3	3	3	0	Poor performance
(2)	Interstate	EVL 008		88			88																	Poor performance and false acceptance
(3)	Super Soft, Inc.	Scratch pad with Voice Drive + TECMAR voice recognition board													5	5	4	3	4	3	5	4	0	Software for use with IBM-PC + 128K Ram + TECMAR voice recognition board. Price includes TECMAR board
(3)	Super Soft, Inc.	Scratch pad with Voice Drive													5	5	4	3	4	3	5	4	0	Software for use with Texas Instruments PC and 128K RAM and TI voice board
(4)	VOTAN	V5000A													3	5	4	4	2	3	3	3	0	Digital voice record and playback

Figure 2.1.2 Results of Voice Recognition Survey

VOICE RECOGNITION SYSTEMS

FEATURES

(Survey Response)	MANUFACTURER, LOCATION	DEVICE(S)	PRICE $K	SYNTHESIS OPTION	MODE: CHIP	MODE: BOARD	MODE: SYSTEM	MODE: SOFTWARE	AVAILABILITY: RUGGEDIZED	AVAILABILITY: MILITARY QUALIFIED	AVAILABILITY: MILITARY QUALITY	AVAILABILITY: COMMERCIALLY AVAILABLE	AVAILABILITY: PROTOTYPE	AVAILABILITY: DESIGN	AVAILABILITY: SUPPORT SOFTWARE	TYPE: DISCRETE WORD	TYPE: DISCRETE WORD, RAPID INPUT	TYPE: CONNECTED WORD	TYPE: CONTINUOUS SPEECH	TYPE: SELECTED WORD IN CONTEXT	TYPE: SYNTAX CAPABILITY	USER INTERFACE: SPEAKER DEPENDENT	USER INTERFACE: SPEAKER INDEPENDENT	USER INTERFACE: SPKR VERIFICATION SECURITY	NO. OF WORDS: BASIC NUMBER	NO. OF WORDS: OPTION NUMBER	UTTERANCE LENGTH (SECONDS)
(4)	VOTAN	V6040	8				✓					✓				✓					✓	✓	✓		256		1
(4)	Interstate	SYS 300	5				✓					✓				✓					✓	✓			200		1.25
(5)	Interstate +	VRT-101 (5 terminals) and link to Univax	40K			✓		✓				✓				✓						✓			65	100	1
(6)	Threshold	15	12.5				✓					✓				✓						✓			48		.5
(6)	Threshold	5					✓					✓				✓						✓			100		.55
(6)	Interstate	12	3.0				✓					✓				✓						✓			55		.3
(7)	Verbex		17.9	✓			✓					✓							✓		✓				120	360	100+
(11)	Infovox	RA-101	3.0			✓	✓					✓				✓						✓			45	3500	2.5

PERFORMANCE BASED ON EXISTING OR READILY AVAILABLE DATA

NOISE EFFECTS (ENTER NO.) — USE/SUSCEPTIBILITY — ENTER RATING: 1 POOR THRU 5 GOOD

MANUFACTURER, LOCATION	DEVICE(S)	ACCURACY WITH DIRECTIONAL MICROPHONE: % STD TAPE TEST	% QUIET ROOM (55 db)	% MAX Db FOR SAME	ACCURACY WITH NOISE CANCELLING MICROPHONE: % STD TAPE TEST	% QUIET ROOM (55 db)	% MAX Db FOR SAME	(OTHER): % STD TAPE TEST	% QUIET ROOM (55 db)	% MAX Db FOR SAME	ACCURACY IN %: QUIET AREA	IN USE AREA	TRAIN IN NOISE	EASE OF ADPT./IMPLMTN	TRAINING EASE	PERF CONSISTENCY	AMBIENT NOISE PERF	PULSE NOISE PERF	MIKE DISTANCE	USER ADAPTATION	VOICE VAR (DAILY, COLD ETC.)	LANGUAGE DEPENDENT (YES ✓ NO 0)	COMMENT (INCLUDES WHO'S CHIP, SPECIAL EQUIPMENT SIZE/VOLUME, WEIGHT, METHOD, WHO'S COMPUTER, ETC.) & FIVE YEAR FORECAST (SYSTEMS CAPABILITIES, ETC.)
VOTAN	V6040													3	5	4	4	2	3	3	3	0	Digital voice record and playback
Interstate	SYS 300													4	4	4	3	2	3	3	3	0	
Interstate +	VRT-101 (5 terminals) and link to Univax				98	98	75							4	4	1	3	4	4	2	4	0	This is not an aircraft application. It is used in 55-80dB noise environment, near runway. No problem
Threshold	15						95				98	95	96	4	3	4	4	2	3	3	3	0	Has been in use 4 years
Threshold	5	✓					96				98			3	3	3	3	2	3	4	4	0	Has been in use 4 years
Interstate	12						95				97			4	4	3	4	3	3	3	4	0	Has been in use 2 years
Verbex					99.2		85							4	3	5	5	4	4	5	4	0	Special computer designed for continuous speech recognition
Infovox	RA-101							TI 99.3						4	4	4	3		4			0	Isolated bench mark used in Ericsson intercom system (proceedings in VDESC in Chicago, 1983)

Figure 2.1.2 Results of Voice Recognition Survey (Continued)

VOICE RECOGNITION SYSTEMS

FEATURES

(Survey Response) MANUFACTURER, LOCATION	DEVICE(S)	PRICE $K	Synthesis option	MODE (enter ✓): Chip	Board	System	Software	AVAILABILITY (enter ✓): Ruggedized	Military qualified	Military quality	Commercially available	Prototype	Design	Support software	TYPE (enter ✓): Discrete word	Discrete word, rapid input	Connected word	Continuous speech	Selected word in context	Syntax capability	USER INTERFACE (enter ✓): Speaker dependent	Speaker independent	Spkr verification, security	NO. OF WORDS (enter no.): Basic number	Option number	Utterance length (seconds)
(11) Threshold Technology	Auricle 1	1.2			✓	✓					✓				✓						✓			40		1.2
(12) Threshold Technology	Model 680	3				✓					✓				✓						✓			120		1
(12) Interstate	ATO	3				✓					✓				✓						✓			120		1
(12) NEC America	SR100	2				✓					✓				✓			✓			✓			120		5
(12) Bell Labs		5				✓					✓							✓		✓				10K		7
(14) HYCOM, Inc.	ID-150 (1085)	2-3		✓ (1085)	✓	✓	✓					✓			✓						✓		✓	50	X (Infinite off-board capacity)	1.2
(15) Interstate	VRT-101	5				✓	✓				✓				✓						✓			100		1.25
(15) Interstate	VRM-102	2			✓						✓				✓						✓			100		1.25

PERFORMANCE BASED ON EXISTING OR READILY AVAILABLE DATA

MANUFACTURER, LOCATION	DEVICE(S)	NOISE EFFECTS (enter no.): Accuracy with directional microphone: % std tape test	% quiet room (55 db)	% max Db for same	Accuracy with noise cancelling microphone: % std tape test	% quiet room (55 db)	% max Db for same	(Other): % std tape test	% quiet room (55 db)	% max Db for same	USE/SUSCEPTIBILITY: Accuracy in %: Quiet area	In use area	Train in noise	Enter rating 1 poor thru 5 good: Ease of adpt/implmtn	Training ease	Perf consistency	Ambient noise perf	Pulse noise perf	Mike distance	User adaptation	Voice var (daily cold etc.)	Language dependent (yes ✓ no 0)	COMMENT (includes who's chip, special equipment, size/volume, weight, method, who's computer, etc.) & FIVE YEAR FORECAST (systems capabilities, etc.)
(11) Threshold Technology	Auricle 1										93 with digits			3	3	4	4		2			0	
(12) Threshold Technology	Model 680					60					60	60	50	2	3	2	2	3	4	2	1	0	
(12) Interstate	ATO					90					90	80	80	1	3	3	1	3	4	3	1	0	
(12) NEC America	SR100					95					95	95	95	3	5	5	5	3	4	4	3	0	
(12) Bell Labs					99	99	99							5	5	5	5	5	5	5	5	✓	By year 1989!
(14) HYCOM, Inc.	ID-150 (1085)					98.5	90				98.5			5	5	5	5	5	4	5	5		LPC and dynamic programming 12 in x 12 in PC board with multi-bus compatible pins
(15) Interstate	VRT-101													5	3	3						0	
(15) Interstate	VRM-102													2	3	3						0	

Figure 2.1.2 Results of Voice Recognition Survey (Continued)

VOICE RECOGNITION SYSTEMS

FEATURES

(Survey Response) MANUFACTURER, LOCATION	DEVICE(S)	PRICE $K	SYNTHESIS OPTION	MODE (ENTER ✓): CHIP	BOARD	SYSTEM	SOFTWARE	AVAILABILITY (ENTER ✓): RUGGEDIZED	MILITARY QUALIFIED	MILITARY QUALITY	COMMERCIALLY AVAILABLE	PROTOTYPE	DESIGN	SUPPORT SOFTWARE	TYPE (ENTER ✓): DISCRETE WORD	DISCRETE WORD, RAPID INPUT	CONNECTED WORD	CONTINUOUS SPEECH	SELECTED WORD IN CONTEXT	SYNTAX CAPABILITY	USER INTERFACE (ENTER ✓): SPEAKER DEPENDENT	SPEAKER INDEPENDENT	SPKR VERIFICATION SECURITY	NO. OF WORDS (ENTER NO.): BASIC NUMBER	OPTION NUMBER	UTTERANCE LENGTH (SECONDS)
(18) VOTAN	V5000	6.3	✓		✓	✓	✓	✓	✓	✓	✓			✓	✓	✓				✓	✓			225	—	2
(18) VOTAN	V6040	8.5	✓		✓	✓	✓	✓	✓	✓	✓			✓	✓	✓				✓	✓	✓		225	—	2
(18) Scott Instrument	Shadow VET-2	.795			✓		✓				✓				✓	✓					✓			40	—	1.6
(18) Voice Machine Communications	Voice input module	.795			✓		✓				✓				✓	✓					✓			80	—	1.6

?

SIR = Speaker Independent Recognition

PERFORMANCE BASED ON EXISTING OR READILY AVAILABLE DATA

MANUFACTURER	DEVICE(S)	NOISE EFFECTS (ENTER NO.): ACCURACY WITH DIRECTIONAL MICROPHONE: % STD TAPE TEST	% QUIET ROOM (55 db)	% MAX Db FOR SAME	ACCURACY WITH NOISE CANCELLING MICROPHONE: % STD TAPE TEST	% QUIET ROOM (55 db)	% MAX Db FOR SAME	(OTHER): % STD TAPE TEST	% QUIET ROOM (55 db)	% MAX Db FOR SAME	USE/SUSCEPTIBILITY: ACCURACY IN %: QUIET AREA	IN USE AREA	TRAIN IN NOISE
(18) VOTAN	V5000	—	99+	80	—	99+	100	—	—	—	99+	99+	99+
(18) VOTAN	V6040	—	99+	80	—	99+	100	—	—	—	99+	99+	99+
(18) Scott Instrument	Shadow VET-2	—	—	—	—	98	55	—	—	—	98	93	85
(18) Voice Machine Communications	Voice input module	—	—	—	—	98	55	—	—	—	98	93	85

MANUFACTURER	DEVICE(S)	ENTER RATING 1 POOR THRU 5 GOOD: EASE OF ADPT./IMPLMTN	TRAINING EASE	PERF. CONSISTENCY	AMBIENT NOISE PERF	PULSE NOISE PERF	MIKE DISTANCE	USER ADAPTATION	VOICE VAR (DAILY, COLD ETC.)	LANGUAGE DEPENDENT (YES ✓ NO 0)	COMMENT (INCLUDES WHO'S CHIP, SPECIAL EQUIPMENT, SIZE/VOLUME, WEIGHT, METHOD, WHO'S COMPUTER ETC.) & FIVE YEAR FORECAST (SYSTEMS, CAPABILITIES, ETC.)
(18) VOTAN	V5000	4	5	5	5	5	5	5	5	0	> Also has excellent quality voice response at low bit rates > RS-232 or multi-bus I/O
(18) VOTAN	V6040	4	5	5	5	5	5	5	5	0	> SIR is language dependent. vocabulary = "0-9 + yes + no" Forecast = speaker verification > and continuous speech in 1984
(18) Scott Instrument	Shadow VET-2	5	5	4	3	1	1	5	4	✓	> Phoneme-based for Apple or Tandy TRS80 computers > Forecast = SIR at chip set level
(18) Voice Machine Communications	Voice input module	4	4	4	3	2	1	5	4	✓	> Phoneme-based for Apple range computers

Scott and VMC are language dependent. restricted to European languages

Figure 2.1.2 Results of Voice Recognition Survey (Continued)

VOICE SYNTHESIS SYSTEMS

FEATURES

(SURVEY RESPONSE) MANUFACTURER, LOCATION	DEVICE(S)	PRICE $K	RECOGNITION OPTION	MODE (ENTER ✓): CHIP	BOARD	SYSTEM	SOFTWARE	AVAILABILITY (ENTER ✓): RUGGEDIZED	MILITARY QUALIFIED	MILITARY QUALITY	COMMERCIALLY AVAILABLE	PROTOTYPE	DESIGN	SUPPORT SOFTWARE	TYPE (ENTER ✓): STORED WORDS PHRASES	DIGITIZED CONDENSED	PHONEME CODING SYNTHESIS	TEXT TO SPEECH
(1) Speech Plus	Prose 2000	6.0				X					X							X
(8) Street Electronics	Echo Speech Prod.	0.2				X								X	X	X	X	X
															?			
(9) General Digital Corporation	GDX-Speech-TI	0.285			X						X			X		X		
(10) Microvoice Systems Corporation	Voiceboard	0.3			X	X	X			X	X	X	X	X	X			
(10) Microvoice Systems Corporation	Microsound	1.2			X	X	X			X	X	X	X	X		X		
(11) Infovox	SA 101	3.0			X						X							X
(11) Texas Instrument	TMS 5100-5220	—		X							X				X			
(11) General Instruments	0256	—		X							X				X			

PERFORMANCE BASED ON EXISTING OR READILY AVAILABLE DATA

(SURVEY RESPONSE) MANUFACTURER, LOCATION	NO. OF WORDS & RATE (ENTER NO.): BASIC NUMBER	OPTION NUMBER	SECONDS OF SPEECH STORAGE	BITS PER SECOND OF SPEECH	MAXIMUM SUSTAINED RATE WORDS PER MINUTE	USE QUALITY (ENTER ✓): NATURAL SOUND	SYNTHETIC SOUND		ENTER RATING 1 POOR THRU 5 GOOD: INTELLIGIBILITY	VOICE QUALITY RATING	EASE OF ADPT IMPLMTN	LANGUAGE DEPENDENT (YES ✓ NO 0)	COMMENT (INCLUDES WHO'S CHIP SPECIAL EQUIPMENT SIZE/VOLUME WEIGHT METHOD, WHO'S COMPUTER, ETC.) & FIVE YEAR FORECAST (SYSTEMS CAPABILITIES ETC.)
(1) Speech Plus							X		3	2		✓	Best text-to-speech to date
(8) Street Electronics	Unlimited	Unlimited	Unlimited	9600	Unlimited	X	X		5	5	5	0	
(9) General Digital Corporation	206	~206			~60	X			5	3	5	✓	Multibus expansion module,
			106	1200									based on TMS 5220, comes with
													TMS 61002 industrial vocabulary
													and socket for optional vocabulary
													Unlimited speech via LPC-10
(10) Microvoice Systems Corporation (Voiceboard)	VB 2000	A,B,C	480	1200	?	X			4	4	5	0	TI (Texas Instruments) LPC
(10) Microvoice Systems Corporation (Microsound)	MS 1000	A,B,C	120	9600	?	X			5	5	5	0	Proprietary modified PCM
(11) Infovox					250		X		4	3	5	✓	Six languages now
													Future: high naturalness, less-robotic voice, several voices, miniaturized version with more languages
(11) Texas Instrument	100			1600				X	4	3	2	✓	
								Neither					
(11) General Instruments	100			2000				X	4	4	2	✓	

Figure 2.1.3 Results of Voice Synthesis Survey

VOICE SYNTHESIS SYSTEMS

(Survey Response)

MANUFACTURER, LOCATION	DEVICE(S)	PRICE $K	FEATURES: RECOGNITION OPTION	MODE (ENTER ✓): CHIP	MODE: BOARD	MODE: SYSTEM	MODE: SOFTWARE	AVAILABILITY (ENTER ✓): RUGGEDIZED	AVAILABILITY: MILITARY QUALIFIED	AVAILABILITY: MILITARY QUALITY	AVAILABILITY: COMMERCIALLY AVAILABLE	AVAILABILITY: PROTOTYPE	AVAILABILITY: DESIGN	AVAILABILITY: SUPPORT SOFTWARE	TYPE (ENTER ✓): STORED WORDS PHRASES	TYPE: DIGITIZED CONDENSED	TYPE: PHONEME CODING SYNTHESIS	TYPE: TEXT TO SPEECH	PERFORMANCE (BASED ON EXISTING OR READILY AVAILABLE DATA) — NO. OF WORDS & RATE (ENTER NO.): BASIC NUMBER	OPTION NUMBER	SECONDS OF SPEECH STORAGE	BITS PER SECOND OF SPEECH	MAXIMUM SUSTAINED RATE WORDS PER MINUTE	USE QUALITY (ENTER ✓): NATURAL SOUND	SYNTHETIC SOUND	USE QUALITY (ENTER RATING 1 POOR THRU 5 GOOD): INTELLIGIBILITY	VOICE QUALITY RATING	EASE OF ADPT IMPLMTN	LANGUAGE DEPENDENT (YES ✓ NO 0)	COMMENT (INCLUDES WHO'S CHIP, SPECIAL EQUIPMENT, SIZE, VOLUME, WEIGHT, METHOD, WHO'S COMPUTER, ETC.) & FIVE YEAR FORECAST (SYSTEMS CAPABILITIES, ETC.)
(12) Votrax Company	TNT	0.4				X					X							X					120		X	1	1	4	✓	Use with Apple computer
(12) Texas Instruments	TMS 5200				X						X				X									X						
(12) Speech Plus	PR2000	3.5			X						X							X					200		X	4		4	✓	Use with HP-16 computer
(12) Speech Plus	PR2020	4.8				X					X							X					200		X	4		4	✓	Use with HP-16 computer
(12) Texas Instruments		2.0			X						X							X					200	X		5	5	5	0	By year 1989
(13) Intex Micro Systems	Text-to-speech synthesizer	0.4			X						X				X		X	X				70-100 ?	2		X	4	3	4	✓	> Votrax SC-01 plus auxiliary circuity > 5 in x 7 in x 3 in > RS-232 and parallel
(13) Intex Micro Systems	Speech Synthesizer	1.0			X							X			X	X			Optional	12K-32K			2	X		5	5	5	0	> ADPCM
(13) Intex Micro Systems	Speech Synthesizer	0.7			X							X			X		X	X	Optional	100-200			10-12	X		5	5	4	✓	Silicon System SSI-263 chip

Figure 2.1.3 Results of Voice Synthesis Survey (Continued)

VOICE SYNTHESIS SYSTEMS

(Survey Response)

MANUFACTURER, LOCATION	DEVICE(S)	PRICE $K	RECOGNITION OPTION	MODE (ENTER ✓): CHIP	BOARD	SYSTEM	SOFTWARE	AVAILABILITY (ENTER ✓): RUGGEDIZED	MILITARY QUALIFIED	MILITARY QUALITY	COMMERCIALLY AVAILABLE	PROTOTYPE	DESIGN	SUPPORT SOFTWARE	TYPE (ENTER ✓): STORED WORDS PHRASES	DIGITIZED CONDENSED	PHONEME CODING SYNTHESIS	TEXT TO SPEECH	NO. OF WORDS & RATE (ENTER NO): BASIC NUMBER	OPTION NUMBER	SECONDS OF SPEECH STORAGE	BITS PER SECOND OF SPEECH	MAXIMUM SUSTAINED RATE WORDS PER MINUTE	USE QUALITY (ENTER ✓): NATURAL SOUND	SYNTHETIC SOUND	(ENTER RATING 1 POOR THRU 5 GOOD): INTELLIGIBILITY	VOICE QUALITY RATING	EASE OF ADPT. IMPLMTN	LANGUAGE DEPENDENT (YES ✓, NO 0)	COMMENT (INCLUDES WHO'S CHIP SPECIAL EQUIPMENT SIZE VOLUME WEIGHT METHOD WHO'S COMPUTER ETC) & FIVE YEAR FORECAST (SYSTEMS CAPABILITIES ETC)
(14) Hycom, Inc.		0.5	X	X	X	X						X				X						2400	Real ~		X	5	5	5	✓	> LPC plus proprietary pitch and
				Chip Set								Est 4/84											Time							voicing algorithm
(16) Micro Mint	Microvox	0.299				X					X			X	X		X	X			3000				X	5	5	5	0	> Votrax SC-01A chip
																					Bytes									> RS-232 interface
																														> Text-to-speech algorithm
																														> 64 levels of inflection
																														> 64 phonemes
(16) Micro Mint	Sweetalker	0.150			X						X			X			X	X						X		5	5	5	0	> Votrax SC-02 chip
	II																													> 256 phonemes
																														> 24 levels of inflection
																														> Text-to-speech algorithm
(17) Speech Plus	Speech	1.2			X						X				X				—	600	360	2200	N/A	X		5	5	5	0	> Also SP1100 downloadable
	1000																													available
	(Multibus)																													> SP1020 packaged board
(17) Speech Plus	Prose	3.0			X						X							X	N/A	N/A	N/A	~150	275	X		4	4	5	✓	> Also PR2020 packaged board
	2000																													
(17) Speech Plus	Call Text																					~150	250							→ Also includes telephone interface
	5000	2.7			X						X			X				X				~150	250		X	4	4	5	✓	← IBM-PC compatible board
	5050	3.0				X					X			X				X				~150	250		X	4	4	5	✓	← RS-232 peripheral (to 6 channels)
	5100	9.8				X					X			X				X							X	4	4	5	✓	← programmable multichannel system
(17) Speech Plus	Speech	0.1					X				X				X	X						2400			X	4	4	3	0	> Encoding services for various
	Encoding	per																				Typical								single-chip synthesizers
		Word																												
(18) VOTAN	V5000	6.3	X		X	X		X	X	X	X			X	X	X			255	•Unlimited when		Ranges		X		5	5	5	0	•This is not really voice synthesis, but
	V6040	8.5	X		X	X														connected to		from 4000 to								digitized recordings, compressed and
								X	X	X	X			X	X	X			255	host computer		14.400								stored for playback under host
																														control
																														•Messages can be easily changed

PERFORMANCE BASED ON EXISTING OR READILY AVAILABLE DATA

Figure 2.1.3 Results of Voice Synthesis Survey (Continued)

Voice Recognition Companies/Systems Reported on by Reviewers:

- Bell Laboratories (Model x)
- Hycom (Prototype Model ID-150)
- Infovox, Inc. (Model RA-101)
- Interstate Electronics, Inc. (Models VRT-101, EVL-008, Sys-300,VRM-102, ATD)
- NEC America, Inc. (Models DP-200 and SR-100)
- Scott Instruments, Inc. (Model Shadow VET-2)
- Supersoft, Inc. (Software for T.I. and TECMAR boards)
- Threshold (Models 5, 15, 680, and Auricle 1)
- Verbex, Inc. (Model 3000)
- Voice Machine Communications (Model Voice Input Module)
- Votan, Inc. (Models V5000 and V6040)

Voice Synthesis Companies/Systems Reported on by Reviewers:

- General Digital Corp. (Model GDX-SPEEDH-TI)
- General Instruments
- Hycom, Inc. (Prototype System)
- Infovox, Inc. (Model SA101)
- Intex Micro Systems (Model Text-to-Speech Synthesizer and Speech Synthesizer)
- Micro Mint, Inc. (Models Microvox, Sweetalker, and Microvox II)
- Microvoice Systems Corp. (Models Voiceboard and Micro Sound)
- Speech Plus, Inc. (Models PROSE-200, PR 2020, Speech 1000,...)
- Street Electronics (Model Echo Speech)
- Texas Instruments, Inc. (Models TMS-5100, TMS-5220, TMS-5200)
- Votan, Inc. (Models V5000 and V6040)
- Votrax Co. (Models TNT, SC-01 and SC-02)

Figure 2.1.4 Systems Identified by Letter Survey

In response to the survey, Dr. David S. Pallett of the National Bureau of Standards, Gaithersburg, Maryland, sent a draft copy of "Guidelines for Performance Assessment of Speech Recognizers." The guidelines are being drawn by Dr. Pallett and a group of eight industry and government voice recognition specialists.

2.1.2.3 Survey of Voice Research and Application Centers

Additionally, five leading U.S. research and application centers were surveyed based on their experience with voice recognition and synthesis equipment:

- Army Avionics Research and Development Activity (AVRADA), Ft. Monmouth, New Jersey
- Air Force Flight Dynamics Laboratory (AFFDL), Wright-Patterson AFB, Dayton, Ohio
- Massachusetts Institute of Technology (MIT), Department of Electrical Engineering and Computer Science, Cambridge, Massachusetts
- Naval Air Development Center (NADC), Warminster, Pennsylvania
- Rome Air Development Center (RADC), Rome, New York

All had experience with two or more voice manufacturers' equipment and added to the data base on the strengths and weaknesses of each system. Their independent research and applications experience provided important insight to the actual state of the art in voice technology, how well it actually performs in the field (specifically in aircraft), and what can be expected from voice systems in the near future (next five years).

AVRADA has built a special sound chamber for testing voice recognition and synthesis systems in ambient sound levels similar to Army helicopters. Systems tested include IEC Voterm, Votan, and Intel voice recognition systems. Votrax and Intex Talker voice synthesis are also used. AVRADA plans in the near future to evaluate flight and military quality voice interactive systems from International Telephone and Telegraph (ITT), Lear Siegler, Inc. (LSI), and Texas Instruments (TI). These will be tested in a UH-60 helicopter simulator and then on a UH-60 test aircraft.

AFFDL emphasis is on voice systems as part of the AFTI F-16 program (in collaboration with NADC and NASA). Testing and evaluation of five voice systems is complete and one is being selected for the Phase II flight test. The five systems are Couzet, Lear Siegler, ITT, SCI, and TI. Since the selection process was in progress, only general information about the tests could be discussed.

AFFDL, in conjunction with the Aerospace Medical Research Laboratory (AMRL), at Wright-Patterson AFB, has created a systematic voice test data base (on audio tape) for testing voice recognition systems. The data base included six subjects (five F-16 pilots) speaking a vocabulary of 70 words and had two parts. The first part was for training the recognition systems with each word repeated five times. The second part was used for testing and the words were spoken in a manner as they would be in an actual flight test. This data base was used for the AFTI Phase II testing and each system was tested under the same conditions and with the same vocabulary. This data base has been made available to anyone interested in testing a voice system.

MIT discussions were application-oriented, covering work at MIT and research for the Defense Advanced Research Projects Agency (DARPA). Key areas of interest were developing capability for independent speech recognition systems and basic studies of the characterization of speech. The near-term problems under investigation include:

- How to get large vocabularies with isolated word recognition systems
- How to get small vocabularies with connected word recognition systems
- How to use large vocabularies with recognition system without each user having to train on all words
- How to extract specific linguistic features from several phrases so that a recognition system can generalize the features to other words in its vocabulary
- How the environment and stress can corrupt speech

NADC is acquiring a Texas Instruments Professional Computer with Voice Command Option (recognition and synthesis capabilities) and a military quality voice system from LSI. They have worked with Votan and Interstate Electronics Inc. (IEC) voice recognition products. The LSI system will be adapted to a Navy F-18 fighter aircraft to test and evaluate voice applications in Navy aircraft. Flight tests of the voice system in the F-18 are scheduled for August 1984. The approach will be similar to that applied for the F-16 Advanced Fighter Technology Integrator (AFTI) program sponsored by Air Force, Navy, and NASA. Initial voice applications in the F-18 include navigation support (entering waypoints), fuel status, interactive checklists (takeoff and landing), and weapon programming. The voice system will be used to directly input data, such as navigational waypoints and therefore bypass the extensive switch activations currently required of F-18 pilots. The TI voice system will be used to investigate other potential voice applications in Navy aircraft. One of the first projects was to look at how an interactive voice system could improve efficiency of P-3 monitoring stations.

RADC's voice work was the most extensive of the three military centers visited and is concerned with voice verification, enhancement (to distinguish speech from background noise), and recognition for Air Force surveillance requirements. Their research and testing are done with equipment built by IEC, Votan, TI, and ITT, in addition to equipment designed and built inhouse. These are two key points of interest on work being done at RADC. First is that they are developing voice recognition algorithms that will work within a wider ambient noise and pilot stress envelope. Second, they have developed a voice enhancement system that does an impressive job of improving voice signal to noise ratios in real time. This system was demonstrated as a front end to a voice recognition system and converted unusable input signals to accurately interpreted ones.

From this survey of selected technology research and development centers, subjects of importance to this study are covered in the following discussion.

Flight/military quality voice recognition and synthesis systems for high gravity (g_n) and high noise environments are both technically feasible and available today. In order to use them, however, a number of factors must be taken into account, including need to operate in isolated word recognition mode, need to limit vocabulary size, and need to eliminate acoustically similar words. Under these conditions several systems have been designed to work adequately in high ambient noise-high acceleration conditions, and to meet military qualification requirements: Crouzet, Lear Seigler, Inc., International Telephone and Telegraph, SCI Systems, and Texas Instruments. With speech enhancement techniques, the performance of these systems in the high ambient noise environment should be further improved. Although the present speech enhancement techniques do not significantly improve the performance of the systems at lower noise environments (below 90 to 95 dB), they could evolve to an ability to enhance selected voice conditions at lower levels. (The ambient noise level in most commercial jet cockpits is below 90 to 95 dB and in late models between 70 dB and 80 dB.

However, a concern that was repeated several times during the trip was that voice systems should not be "dropped" into the cockpit. The addition of voice should be part of an intelligent system upgrade where it is integrated with redesigned switches and displays. For example, voice is not efficiently used if it merely mimics switches. One alternative is that voice recognition can be used to directly access an area of interest instead of stepping through a hierarchical paging scheme.

As an extension to the intelligent system upgrade, the incorporation of artificial intelligence (AI) systems will improve the performance and usefulness of voice systems. AI could control voice recognition systems by defining which words should be recognized at any one time, allowing for variable syntaxing, and checking command strings for illogical words. AI could control voice synthesis systems so that pilots could be questioned about potentially dangerous commands or selections. It could also prioritize voice messages or suppress them to avoid noise clutter in the cockpit.

Synopsis: Progress toward future capability now being made by voice recognition and synthesis researchers was particularly noted at the visit to MIT. Many major advancements have been made or are in the works and should be transferred to industry in the next five years. Even in the short time from the survey to preparation of this document, improved recognition accuracy, ability to handle connected speech, and lower system costs have been observed.

Voice recognition systems that are presently on the market use little, if any, knowledge about the individual. They primarily use classical signal pattern recognition techniques. The advantages of this approach are that they are language independent and require only adaptation to the speaker's physiology, background noise, microphone location, environmental factors, etc. The disadvantages include mandatory training for each word and the potential for changes in the speaker's voice to the extent that accuracy suffers.

Technological progress at the moment is such that the voice recognition hardware currently available is more sophisticated than the software being used on them (recognition algorithms). Research is needed to improve the algorithms so that recognition systems performance can be maintained as speech patterns vary from changing physiological conditions and articulation. With connected word recognizers, improved algorithms should account for coarticulation of strings of words. Again, AI may be important to take full advantage of the hardware capability available today.

Two additional, but important, comments from the survey pertain to evaluating voice recognition systems for use in aircraft. First, it is quite important that realistic/natural vocabularies be used. Representative users should be involved early in any given application program to evolve natural procedures and word structure. Second, accuracy considerations should also include substitution and rejection percentages. The advantages of few rejections and high accuracy may be outweighed by relatively high substitution rates.

Voice synthesis technology is far ahead of voice recognition. Commercial quality synthesis systems are available off the shelf with vocabularies greater than 10,000 words, variable speed, variable inflection, variable tone quality, and that are remotely controllable.

2.1.2.4 Military Flight Quality Voice Recognition Systems

Currently five military flight quality voice recognition/synthesis systems are available. Their manufacturers are Crouzet (French), Lear Seigler, Inc., ITT Defense Communications Division, SCI Systems, Inc., and Texas Instruments, Inc. Progress on the last four systems is due, in part, to the impetus provided by the joint Air Force, Navy, NASA Advanced Fighter Technology Integrator (AFTI) Program, coordinated by the Flight Dynamics Laboratory, Wright-Patterson Air Force Base.

The AFTI program is interested in using voice recognition as an option in manual switching methods for command and control in combat aircraft. The voice evaluation program has three phases, of which the first two have been completed. Phase 0 was a laboratory simulation evaluation to determine the viability of voice as an alternate to manual switch control. Phase 1 goals were (1) to determine the effects of the airborne environment on the pilot's voice and on the recognition algorithm performance, (2) to develop a reliable voice command system (VCS), (3) to demonstrate feasibility in airborne applications, and (4) to establish a basis for further functional studies. (Moore and Ruth, 1984.) Ground simulation preceded a limited flight test in the AFTI F-16 test aircraft.

The primary goal of Phase 2 is to optimize the pilot-vehicle interface. An extensive ground simulation was conducted to determine if voice recognition was at an acceptable state of reliability and performance. The five manufacturers noted above provided systems for evaluation. One or two systems will be chosen for extensive flight testing in the AFTI F-16 test aircraft in 1985.

The results of the Phase 2 ground simulation tests are being prepared for publication. Selection of the voice command system supplier will be announced in the near future. Each voice system manufacturer was given minimum system requirements that included:

- Speaker-dependent voice recognition with no more than five training passes required per word and selective retraining and capability
- Isolated word recognition with a goal of limited connected speech capability
- Ability to operate as a system controller and interface with other aircraft systems via a Military Standard 1553 data bus
- Cockpit-located data transfer module
- Total recognition vocabulary of at least 100 words or phrases with at least 20 nodes (vocabulary subdivisions) and 25 words per node

- Recognition response time from end of word or phrase of 0.5 sec maximum

- Total synthesis vocabulary of at least 200 sec with the complete recognition vocabulary

- Overall isolated word recognition accuracy of 90% with a design goal of 95%. This accuracy must be achieved in an environment of 92 dB ambient noise, 3-g_n force, and pilot using a standard oxygen mask

The AFTI Phase 2 voice command system shows the technical level of current flight quality voice recognition and synthesis systems. Figures 2.1.2 through 2.1.4 and Appendix A include commercial grade voice systems that have performance and capabilities superior to the AFTI voice command systems, but they are neither flight quality nor can they operate in the AFTI noise and g_n-force environment. Although some of AFTI's requirements are unnecessary for commercial aircraft, they have established a standard frame of reference from which requirements for future flight quality voice systems can advance.

2.1.2.5 Technology Capabilities and Limitations

The capabilities of currently available voice recognition and synthesis systems have been discussed in previous sections. For the purpose of clarification and reference by subsequent sections, the capabilities and limitations of present and near-term voice technology are summarized here. The capabilities or advantages of a method are identified with a "+" and limitations are identified with "-".

Voice recognition systems vary in operating modes and capabilities and limitations. For comparison purposes, four categories are used below: isolated word, connected word, continuous speech, and military flight quality systems.

Isolated word voice recognition systems

Capabilities

+ The systems report highest performance in isolated word mode, e.g., 95% to >99% claimed by manufacturers.
+ The number of words that is recognizable at one time with near real-time response is over 100 on some systems. Larger active vocabularies are possible but show a noticeable time lag from voice input to acknowledgement.
+ Performance in high ambient noise is better than connected or continuous word systems.
+ Training is faster than for connected and continuous modes due to not having to deal with coarticulation between words.
+ Speaker independence for a limited vocabulary is available on some isolated word systems.
+ Some systems have word-spotting capability. This allows trained words to be interspersed with nontrained words in word strings and still be recognized.
+ Most systems have ability to pass voice reference tables between host systems. Large total vocabularies are achievable this way.
+ Support software usually permits creation of vocabulary trees and ability to jump from node to node on the tree. Each node will have a predefined subset of the total vocabulary active for recognition.

Limitations

- The systems are not natural to use because the user must carefully say each word and allow space between words.
- Limited active vocabularies of 50 to 100 words or phrases exist on most systems.
- To get few substitution errors the rejection threshold must be raised. This in turn decreases overall performance accuracy because more correct words are also rejected.
- Few work well in high ambient noise. Some can work if trained in the noise. Few can be trained in a quite environment and perform effectively in high noise.
- Performance claimed by manufacturers is rarely observed by users due to less ideal operating conditions and environments.
- Most systems require a signal-to-noise ratio of 20 dB for proper operation.

Connected word voice recognition systems

Capabilities

+ The systems are closer to natural speaking conditions than isolated word systems.
+ Some systems have no limit to the number of words that may be strung together.
+ Some systems have word-spotting capability. This allows nontrained words to be ignored when interspersed with trained words in word strings.
+ Most systems have ability to pass voice reference tables between host systems. Large total vocabularies are achievable this way.
+ Support software usually permits creation of vocabulary trees with ability to jump from node to node on the tree. Each node will have a predefined subset of the total vocabulary active for recognition.

Limitations

- Total active vocabularies are limited to 50 to 70 words or phrases.
- Systems usually will not respond until the speaker has stopped speaking or the input buffer is filled.
- Performance degrades in high ambient noise faster than for isolated word recognition systems.
- All systems are speaker dependent and require training.
- Training usually requires repeating each word or phrase two or more times separately before embedding it in a word string.
- If word-spotting is not available then the user must use only trained words or the system may hang up.

Continuous speech voice recognition systems

Capabilities

+ Users may speak in a natural manner.
+ Some systems have no limit to the number of words that may be strung together.

Limitations

- Total active vocabularies are limited to 50 to 70 words or phrases.
- Systems usually will not respond until the speaker has stopped speaking or the input buffer is filled.
- Performance degrades in high ambient noise faster than for isolated word recognition systems.
- All systems are speaker dependent and require training.
- Training usually requires repeating each word or phrase two or more times separately before embedding in a word string. Training is usually more rigorous than with connected word systems.
- If word-spotting is not available, then the user must use only trained words or the system may hang up.

Flight (military) quality voice recognition systems

Capabilities

+ Isolated mode systems have been demonstrated to operate acceptably in high ambient noise and stress environments.
+ Training is done in a quiet, nonstressful environment.

Limitations

- Demonstrated systems have limited vocabulary capabilities.
- All are speaker dependent and require training on each word or phrase.
- Training usually requires each word to be repeated four to five times.
- Prices on these systems are as much as an order of magnitude and more above comparable commercial systems.

Forecasts are that the capabilities and performance of voice recognition systems will be increased by impressive jumps in the next five years. Of particular significance to present purposes is the estimation of available flight quality voice recognition systems in the time period. Prototype military flight quality connected word recognition systems are now being designed and built, and deliveries to military and commercial airplane manufacturers for evaluation are expected by the end of 1984. These systems will have comparable capacity and performance to commercial systems now available plus preprocessors for noise subtraction. The response time of these systems should approach real time.

Improving the accuracy of recognition systems is also of importance here. One concept uses a scoring system that scores performance on all words and chooses the alternative when the first choice doesn't fit the context. Some recognition systems can pass the scores for recognized words and the second best choices to a host system. AI systems will probably be used in the near future to interact with these voice systems. If a word does not fit in the context of the word string, the confidence level for that word can be examined and the applicability of the second best choice can be examined. In this way the overall accuracy of a voice recognizer could be improved. This is now technically feasible, and if it is not being explored now it soon will be.

As mentioned earlier, the real advancement of voice recognition will happen when recognition systems start using knowledge of speech and not just pattern recognition. Work in this area is being performed now at universities (MIT) and major research centers (Bell Laboratories). The transfer of this information to industry is expected to happen during the next five years.

Voice synthesis systems, as noted earlier, are more advanced than voice recognition systems. Voice synthesis is also in greater use today than recognition, for example in toys, vending machines, and aircraft. The two primary methods used are (1) digitized/condensed whole- or partial-word method, and (2) digitized/condensed phoneme method. Note that the former is not synthesis so much as efficient digital recording and playback.

Digitized and condensed whole- or partial-word synthesis/playback method

Capabilities

+ Poor- to high-quality voice reproduction is possible depending on sampling method (e.g., linear predictive coding) and rate (bits per second).
+ Simple chip-sets for standalone systems are available.
+ Some manufacturers have large existing vocabularies available.
+ Desired voice/tone is selectable, not restricted to manufacturer's vocabulary.
+ Features are easy to interface with or design into user's system.
+ Words can be strung together to form messages or sentences.
+ Flight quality systems are available and being used.

Limitations

- User usually must have manufacturer digitized user's vocabulary. This is expensive and requires large lead time.
- Vocabularies are memory limited. Each word requires memory and must be addressed.

Digitized and condensed phoneme synthesis method

Capabilities

+ Poor- to good-quality voice synthesis is possible depending on number of phoneme bases, variable inflection, tone control, speed control, and extent of word definitions.
+ Chip-sets are available for rudimentary voice synthesis. Chip-sets are more precise, and higher quality voice synthesis is expected in the near term. Technology to do so is available but needs market interest.
+ Some commercially available systems have large predefined vocabularies (>10,000 words), good-quality voice synthesis, variable speed, inflection and tone, and the ability to imitate young and old, male and female.
+ If combined with AI, good-quality phoneme systems could be used for near natural conversation.
+ It is technically feasible to ruggedize current top of the line phoneme systems.
+ Near-term improvements will bring natural sounding voice and larger predefined vocabularies.

Limitations

- Even the best phoneme systems still sound synthetic.
- No good-quality phoneme systems have been ruggedized or flight qualified.

2.1.2.6 Voice Recognition Performance Measures

Voice recognition manufacturers and users often quote the accuracy their respective systems have achieved. Unfortunately, complete documentation on the types of tests that were used to determine the accuracies is rarely available. The most common method is to use a tape to train the system, then use the same tape to test recognition accuracy. Other test conditions are also relevant, e.g., ambient noise. Unfortunately, without an understanding of any additional circumstances of the tests, the accuracy figures are not as useful as could be desired.

Added test information is certainly necessary to expand on the scope of meaning for the quoted accuracy. Also, for example, a system may perform with a 97% accuracy with a given vocabulary, but if 2% of its 3% error are of a substitution type, then it will be unacceptable for some critical tasks where it is better to miss than to execute the wrong command. A recognition system's performance and accuracy, i.e., words correctly recognized out of total number of recognizable words, are dependent on:

- Size of the active vocabulary
- Acoustic similarity of words in the vocabulary
- Number and type of vocabulary words spoken for test
- Number and type of nonvocabulary words spoken for rejection/substitution test
- Nature and level of ambient noise
- Type and location of microphone

- Reproducibility of user's voice
- Quality of voice reference data
- Rejection threshold or criteria used by system

As important an element in overall performance as accuracy is knowledge of the type of errors that occur. Errors can generally be grouped into rejection, substitution (misinterpreting one vocabulary word as another), false rejection (deletion, not recognizing a vocabulary word), or false acceptance (insertion, misinterpreting a nonvocabulary word or noise as a valid vocabulary word). Also of interest is a system's ability to reject nonvocabulary words.

Speech recognizer assessment guidelines in the draft provided by the National Bureau of Standards (ref. 14) suggest the following test data summary for standardized documentation:

- Criteria in Setting the Reject Threshold or Reject Criteria

$$\text{Correct recognition rate (percent)} = \frac{\text{Number of correctly recognized words} \times 100}{\text{Number of test words}}$$

$$\text{Substitution error rate (percent)} = \frac{\text{Number of substitution words} \times 100}{\text{Number of test words}}$$

$$\text{False rejection/deletion error rate (percent)} = \frac{\text{Number of deleted words} \times 100}{\text{Number of test words}}$$

$$\text{False acceptance/insertion error rate (percent)} = \frac{\text{Number of inserted words} \times 100}{\text{Number of test words}}$$

$$\text{Rejection (nonvocabulary) rate (percent)} = \frac{\text{Number of rejection responses} \times 100}{\text{Number of test words}}$$

The above information will help in selecting a system, but more information on conditions of the intended application that affect performance is required to plan for a successful voice recognition application. Information to be considered and tested in candidate recognition systems includes ambient noise interference, speech signal quality, syntaxing schemes, and the confusability of the vocabulary. It is recommended that a designer who plans to use a voice recognition system become familiar with both the influence and methods of controlling these performance factors. Comprehensive performance assessment techniques are presented in a 1983 report by Lea and Woodard (ref. 18).

2.2. Task 2: Appraisal of Voice in Control and Information Transfer

Task 1 examined the status of existing and near-term voice recognition and synthesis technology.

The objective of task 2 was to examine a number of voice recognition and synthesis applications possibilities for commercial aircraft flight decks and simulators. Both existing systems and possibilities with newly identified systems were considered. This section presents (a) information on the aircraft subsystems considered, (b) the rating of voice recognition and synthesis applications for management and control of the subsystems for both existing and predicted systems, and (c) information on applicability of voice systems in systems training.

The flight test voice recognition systems, described in Section 2.1.2.3 for the AFTI project, have a limited vocabulary and do not have the capabilities to adequately perform the full scope of tasks. Other systems have the capabilities but could not be appraised to determine if they could be suitable for experimental flight applications in the more benign commercial aircraft environment. Flight quality voice systems with sufficient capabilities should be available in one to three years. Similar commercial systems are available now.

Communications, navigation, and automatic flight control systems are prime candidates for voice recognition applications. Voice recognition also shows a potential for systems management in data entry and programming applications. Voice synthesis is already in use in commercial cockpits for caution-warning alerts. Several other tasks are considered here such as altitude and airspeed callouts, air traffic control message playback, and interactive checklists. However, the use of voice synthesis must be carefully planned to avoid flight deck noise pollution and possibly multiple messages of differing criticality that could be in conflict and, in turn, create dangerous situations.

2.2.1 Aircraft Applications of Voice Systems

All existing pilot flight deck tasks were categorized into five general functions and related to a commercial aircraft's subsystems (fig. 2.2.1). In order to establish a baseline for later ratings, it was assumed that a digital interface existed for installation of voice systems, e. g., a Boeing 767 flight deck. Using the pilot functions of controlling and monitoring the aircraft subsystems as the baseline, candidate voice applications were identified and a list of potential voice tasks was generated. The tasks ranged from selecting a specific communications radio to programming an autopilot. Each subsystem was examined for its specific requirements: vocabulary size and type, minimum acceptable recognition accuracy necessary, and task frequency. This provided the framework within which potential voice applications could be evaluated.

The potential tasks associated with each subsystem were identified and each task was rated for technical feasibility, utility or advantage to the pilot, associated time and accuracy requirements, and hardware adaptability.

2.2.2 Potential for Interfacing Voice Systems With Aircraft Subsystems

2.2.2.1 Interfacing With Existing Aircraft Subsystems

This section presents discussions concerned with potential for voice systems to interface with existing aircraft subsystems and the potential for applications with new subsystems.

All aircraft subsystems that pilots interact with in the cockpit were considered for possible application of voice recognition and/or synthesis. Each of these subsystems has certain operational requirements that must be met by a voice system if it is to be effective and accepted for control and information transfer. The vocabulary associated with each subsystem's tasks may include unique words as well as words common to other tasks and subsystems. Accordingly, words pilots commonly associate with each task are used as the bases for the vocabularies. The tasks associated with each subsystem and projected requirements of vocabulary size, minimum acceptable recognition accuracy, and task frequency are defined below.

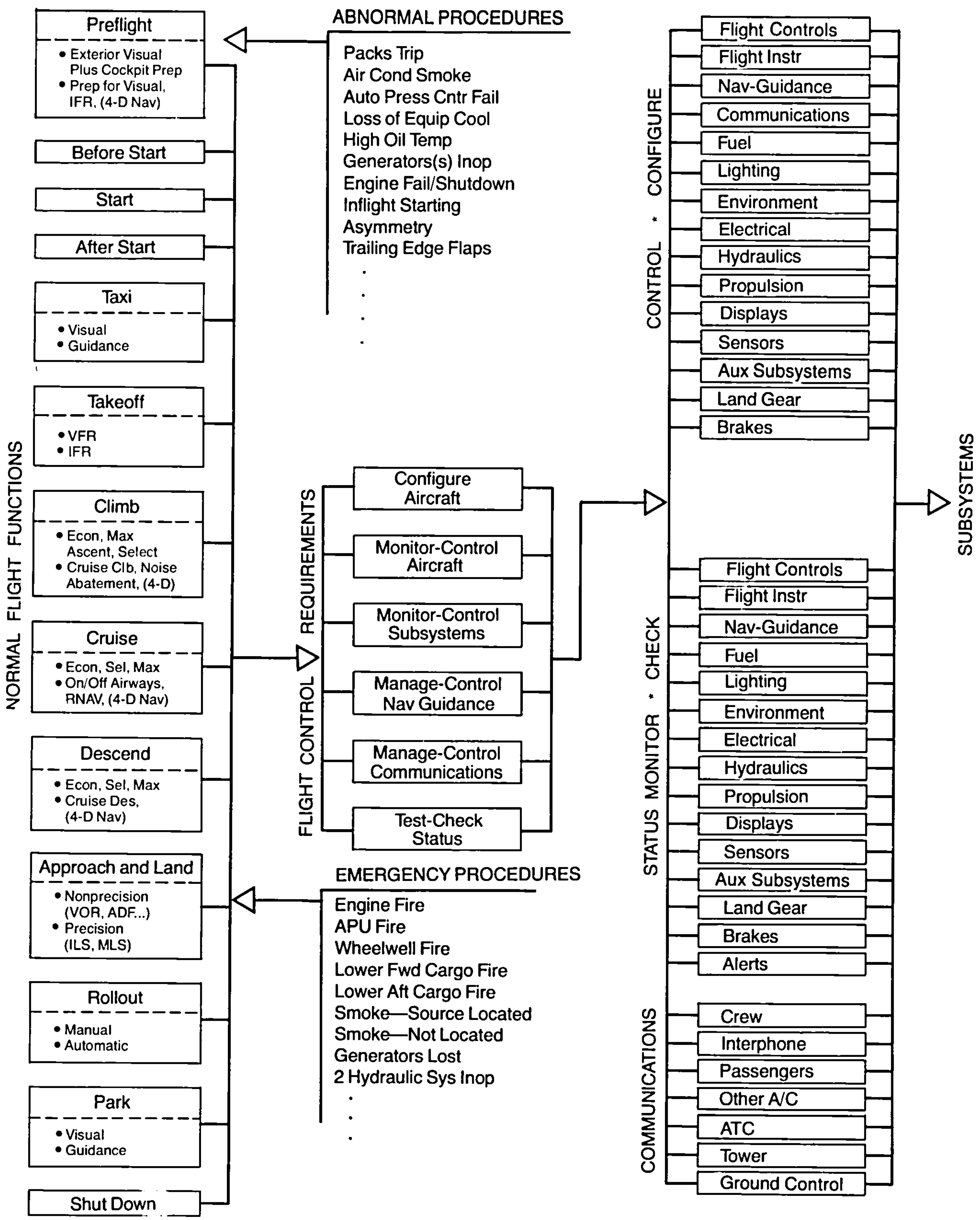
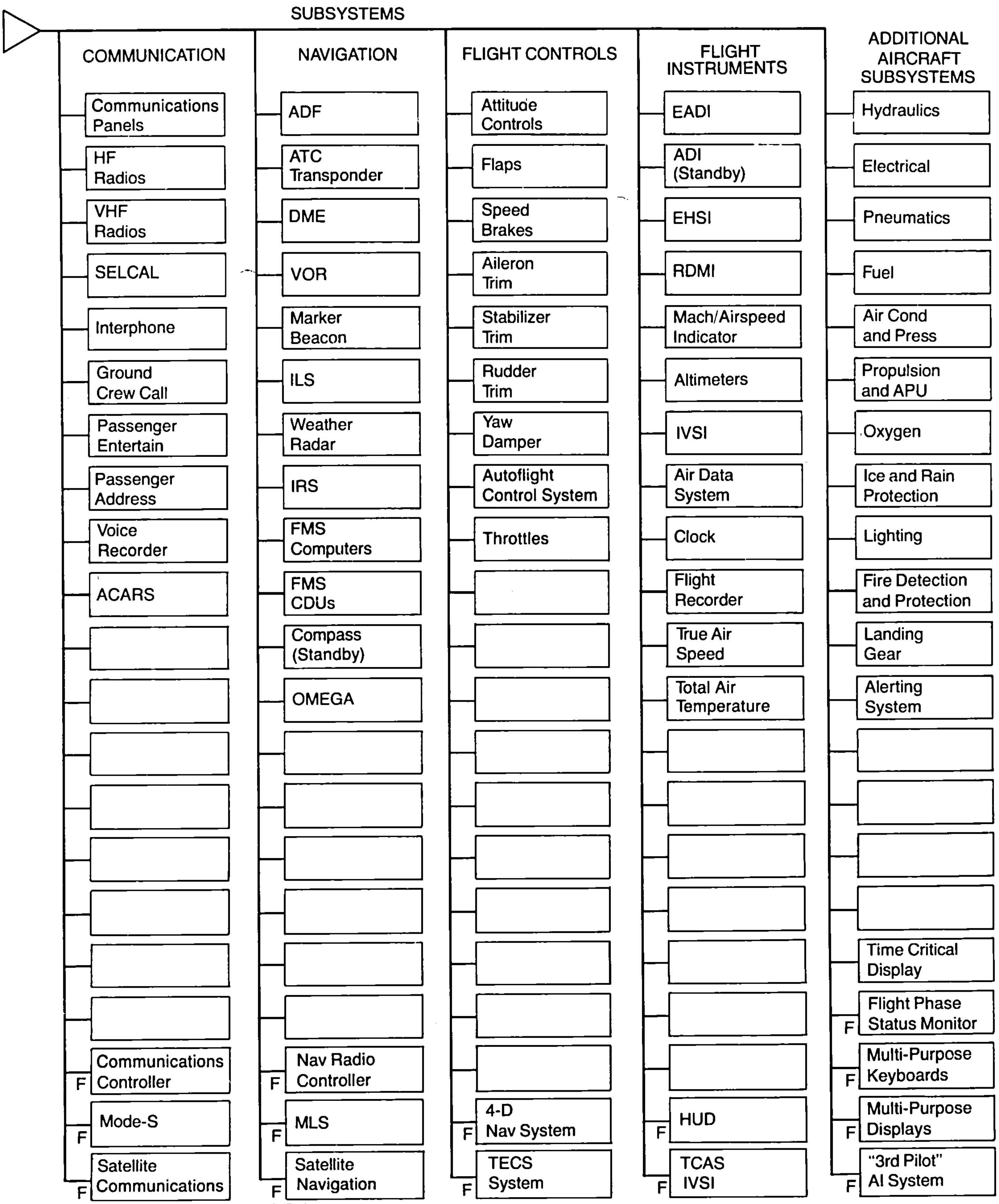


Figure 2.2.1 Aircraft Subsystems Organization Concept



F = Possible Future Aircraft Subsystems

Figure 2.2.1 Aircraft Subsystems Organization Concept (Continued)

Existing communications subsystems, potential tasks, and requirements

Tasks

- Adjust radio volume controls
- Make a radio selection via the communications panel
- Tune radios by identifying the radio and entering frequency
- Tune one or more radios by identifying communications source, e.g., Seattle approach control

Constraints

- Vocabulary size for communications subsystem management should be less than 50 for any one flight. Common words will be 0-9, right, left, center, radio, set, point. Unique words will include VHF, HF, ground control, tower, departure, center, approach, clearance, delivery, departure city, arrival city, and three or four en route cities
- Minimum acceptable recognition accuracy will be ⩾90% overall with acceptable feedback — correction features
- Task frequency is low to medium

Existing navigation subsystems, potential, tasks, and requirements

Tasks

- Select a radio, system, or system mode
- Tune radios by identifying the radio and entering frequency
- Tune one or more radios by identifying signal source
- Enter data into the inertial reference system (IRS), using digits or location identification, probably via control display unit (CDU)

Constraints

- Vocabulary size for navigation subsystem management should be less than 50 words for any one flight. Unique words will include VOR, DME, ILS, ADF, MLS, NAV, navigation, aids, and 10 or more navigation aid identifiers for that flight
- Minimum acceptable recognition accuracy will be ⩾90% overall with acceptable feedback — correction features
- Task frequency is low to medium

Existing flight control subsystems, potential tasks, and requirements

Tasks

- Select positions for flaps, speed brake, and trims
- Select autopilot and fuel management system modes
- Enter data into autopilot and fuel management systems, probably via CDU

Constraints

- Vocabulary size for flight control subsystems should be less than 50 words. Unique words/phrases will include autopilot, autothrottle, mach, knots, speed, altitude, vertical, heading, disconnect, altitude, hold, flight, level, change
- Minimum acceptable recognition accuracy will be ⩾98% overall with acceptable feedback - correction features
- Task frequency is low to continuous

Existing flight instrument subsystems, potential tasks, and requirements

Tasks
- Direct primary attitude controls
- Select speed and height bugs
- Select decision height
- Enter barometric pressure
- Select modes for EADI, EHSI, and EICAS
- Call out airspeed and altitude positions

Constraints
- Vocabulary size for flight instrument subsystems should be less than 50 words for any one flight. Unique words or phrases will include L-NAV, V-NAV, backcourse, localizer, approach, EPR, takeoff, climb, continue, cruise, temp, temperature select, PSI
- Minimum acceptable recognition accuracy will be ⩾90% overall with acceptable feedback - correction features
- Task frequency is low to high

Other existing subsystems, potential tasks, and requirements

Tasks
- Primarily switching and information exchange or transfer (including voice synthesis)

Constraints
- Some subsystems, such as landing gear, are very time critical, and others are rarely used

The minimum acceptable recognition accuracy will vary from task to task and according to different phases of flight. Accuracy is directly tied to the criticality of the task and the time available to verify and possibly correct any errors. If more than one word is required to accomplish an action, then the probability of the correct instruction being understood equals the probability of correctly interpreting each word in the word string. For example, if the individual word accuracy is 95% then the probability of all the words in a five word string being correctly recognized is only 77%! With 98% accuracy for the individual word, the probability of all five correct is 90%. To achieve a probability of 97% for all five, single word accuracy of 99.5% is required. For this study it was assumed that, in general, up to five connected words would be needed to command or enter data to any one system at one time, and AFTI criteria of 90% minimum and 95% desired performance accuracy were adopted as a preliminary frame of reference. It is suspected that 95% to 98% accuracy or better will be necessary to achieve pilot acceptance in a nonexperimental operational context.

A rating scheme was employed to help determine whether voice systems could be used to control or exchange information with the aircraft subsystems mentioned above. The rating scheme used in this section is based one developed by Feuge and Geer (ref. 12). Although the intent of the ratings is much the same, the definitions have been tailored to support the requirements of this study. As part of the rating exercise task, frequency was broken down into several levels from almost nonuse to near continuous use: nonuse = 0 to 1 interactions per hour, low use = 2 to 5, medium use = 6 to 15, high use = 16 to 30, and continuous use = over 30 interactions per hour. A task may be done at a high frequency level during takeoff and be essentially ignored during cruise, as in the case of trim and flaps control.

Each potential voice task is rated on three factors: (1) Technical feasibility: What voice recognition/synthesis capabilities are required for this task? (As a baseline for technical feasibility, the AFTI/F-16 Phase II specification was used to define the flight quality voice recognition/synthesis system that is available today). (2) Utility: Is the application of voice beneficial, equal to, or disadvantageous to crew members performing the task? (3) Time/accuracy requirements: What accuracy is required to do the task due to the timeframe available for correcting errors? In addition to the three user-based rating categories, each potential voice task was assigned a hardware adaptability rating for existing (e.g., B-767) and next generation (year 1990) aircraft. This second rating was to simplify the appraisals by eliminating the constraints of analog-to-digital conversion that would be required for older models; similar applications would be feasible if an analog-to-digital conversion were accomplished - a relatively expensive application to present analog systems in line aircraft. The ratings and definitions for each rating factor are presented on Tables 2.2.1 and 2.2.2.

During the rating process certain assumptions were made and they should be kept in mind when reviewing the scores: (1) pilots using the voice systems would be properly trained with them; (2) pilots would be supportive of the voice system, i.e., would not try to trick the voice recognition system; (3) pilots would use a push-to-talk switch to activate the voice recognition system; (4) a prominently located preentry display would be used to verify the recognition system's interpretation; and (5) each pilot would have a voice recognition system that would default to controlling the speaking pilot's systems but could, if directed, control the other pilot's systems.

The rankings of the potential voice tasks shown in Tables 2.2.1 and 2.2.2 are a composite score based on input from seven Boeing personnel; two were commercial aircraft training pilots, three were flight deck research engineers (two of whom were military multiengine pilots), and two were human factors specialists with extensive backgrounds in commercial flight deck operations and design. The ranks assigned in Tables 2.2.3 and 2.3.3 through 2.3.5 were also based on their comments.

The pilots were not familiar with flight deck voice systems other than conventional aural alerts; therefore, a briefing was given to them on current voice systems capabilities and research applications. The applications noted in the tables above were not ranked by the pilots. Rather, they were questioned about various voice input/output applications on aircraft and simulator flight decks during a several hour session in a B-767 training simulator. Two flight deck research engineers (one who was a military multiengine pilot) sat in the jump seats and discussed with the pilots the use of various flight deck systems and where voice recognition/synthesis might be implemented in order to assess the worth of each application.

The session in the simulator included "flying" a typical commercial flight with a number of abnormal situations added. Several emergency events were also discussed to more fully explore potential voice applications. After the "flight," voice applications with next-generation systems and flight deck simulators were discussed. Throughout the session, the pilots' impressions and responses were recorded and combined afterward to rank the various voice applications.

The flight deck engineers and human factors specialists either specifically ranked each potential voice application or their comments on the applications were gathered during interviews. These engineers and specialists have all been working with voice recognition and synthesis flight deck applications for at least four years. The rankings as listed in Tables 2.2.1 through 2.2.3 and 2.3.3 through 2.3.5 summarize their inputs as well as those of the pilots.

Table 2.2.1 Voice Recognition Ratings for Potential Cockpit Applications

Definitions: Voice <u>Recognition</u> Rating Scheme
Feasibility 1 = FEASIBLE NOW (speaker dependent, isolated word recognition, 70 word total vocabulary, but only 25 available at one time. AFTI/F-16 type) 2 = FEASIBLE IN 1-3 YEARS (speaker dependent, connected recognition, 300 or more words in vocabulary, and 70 available at one time) 3 = FEASIBLE IN 3-5 YEARS (limited training, continuous recognition, >1000 words in vocabulary and 100 to 200 available at one time) 4 = NOT FEASIBLE IN <5 YEARS (limited training,continuous recognition, >5000 words in vocabulary and >1000 available at one time.
Advantages 1 = ADVANTAGE (voice would be advantage to pilots' task, potential hand/eye overload) 2 = NO ADVANTAGE (voice is no advantage, equivalent to existing method of performing task) 3 = DISADVANTAGE (voice could potentially create a disadvantage to pilots performing task)
Action Criticality 1 = NO PROMPT ACTION REQUIRED (time available to correct errors, at least 95% single word accuracy needed for 77% accuracy on 5 connected words) 2 = PROMPT ACTION REQUIRED (little time to correct errors, at least 98% single word accuracy needed for 90% accuracy on 5 connected words) 3 = IMMEDIATE ACTION REQUIRED (almost no time to correct errors, at least 99.5% single word accuracy needed for 97% accuracy on 5 connected words)
Adaptability 1 = NO PROBLEM (bidirectional data bus available, programming may or may not be required) 2 = SOME DIFFICULTY (unidirectional data bus available/in use, i.e., ARINC-429, voice system could replace or substitute for existing control head or new input/output ports could be added) 3 = MORE DIFFICULT (no data bus exists, but digital electronics are incorporated in particular aircraft system therefore can modify particular aircraft system to add interface to voice system) 4 = VERY DIFFICULT (no data bus exists and no convenient electrical hardware to interface to, not worth trouble)

Table 2.2.1 Voice Recognition Ratings for Potential Cockpit Applications (Continued)

Potential Voice Recognition Application	Technical Feasibility Factor	Utility/ Advantage Factor	Time/Accuracy Requirement Factor	Hardware Adaptability Factor	
Communications					
• Switching and selecting modes (E)	1	2	1	3(E)	3(N)
• Volume control (E)	1	3	1	4(E)	3(N)
• Entering frequencies (E)	1	1	2	2(E)	1(N)
• Radio tuning by location ID (E)	1	1	2	2(E)	1(N)
• Selecting and preparing messages	If 1	Then 2			
for Mode-S type data transmission (N)	If 2	Then 1	1	-	1(N)
Navigation					
• Switching and selecting modes (E)	1	2	1	3(E)	1(N)
• Entering frequencies (E)	1	1	2	2(E)	1(N)
• Radio tuning by location ID (E)	1	1	2	2(E)	1(N)
• Programming CDU [IRS, NAV and	If 1	Then 2			
performance management] (E)	If 2	Then 1	2	2(E)	1(N)
• Programming microwave landing	If 1	Then 2			
system [MLS] (N)	If 2	Then 1	2	-	1(N)
Flight Controls					
• Primary attitude controls (E)	2	3	3	4(E)	1(N)
• Selecting positions for flaps, speed brake and trims (F)	1	3	3	3(E)	1(N)
• Select autopilot and fuel management systems modes (E)	1	2	2	2(E)	1(N)
• Entering data to autopilot and throttle (E)	1	1	2	2(E)	1(N)
• Selecting modes for autopilot and advanced fuel management system (N)	1	2	1	-	1(N)
• Programming 4D navigation system (N)	2	1	2	-	1(N)
Flight Instruments					
• Selecting speed and height bugs (E)	1	1	2	2(E)	1(N)
• Entering barometric pressue (E)	1	1	2	1-4(E)	1(N)
• Selecting modes for EADI, EHSI, EICAS and HUD (E/N)	1	2	1	2(E)	1(N)

Note: (E) = Existing systems task (N) = Next generation systems task

Table 2.2.1 Voice Recognition Ratings for Potential Cockpit Applications (Continued)

Potential Voice Recognition Application	Technical Feasibility Factor	Utility/ Advantage Factor	Time/Accuracy Requirement Factor	Hardware Adaptability Factor	
Additional Aircraft Subsystems [Hydraulics, electrical, pneumatics, fuel, air conditioning, engines, APU, anti-ice, rain protection, fire protection, landing gear, crew alerting]					
• Selecting positions and modes (E)	1-2	2-3	1-3	2-4(E)	1-4(N)
• Integrated systems management (N)	If 1 If 2	Then 2 Then 1	1	-	1(N)
Flight Status Monitor					
• Request schematics and checklists (N)	If 1 If 2	Then 2 Then 1	3	-	1(N)
• Request system status (N)	If 1 If 2	Then 2 Then 1	2	-	1(N)
Programmable Multipurpose Keyboard					
• Paging (N)	If 1 If 2	Then 2 Then 1	2	-	1(N)
• Entering data/programming some A/C subsystems (like CDU) (N)	If 1 If 2	Then 2 Then 1	2	-	1(N)
Multipurpose Displays					
• Paging and format requests (N)	If 1 If 2	Then 2 Then 1	2	-	1(N)
• Request and step-through operations checklists (N)	If 1 If 2	Then 2 Then 1	2	-	1(N)
Artificial Intelligence (AI) System					
• Interaction with AI systems recognition/understanding (N)	2	1	2	-	1(N)

Note: (E) = Existing systems task, (N) = Next generation systems task

Table 2.2.2 Voice Synthesis Ratings for Potential Cockpit Applications

Voice Synthesis Rating Scheme
Feasibility 1 = FEASIBLE NOW (digitally compressed voice recording and vocabulary of 100-200 words) 2 = FEASIBLE IN 1-3 YEARS (phoneme type system 10K word vocabulary) 3 = FEASIBLE IN 3-5 YEARS (phoneme type system, >10K word vocabulary, multiple languages)
Advantages 1 = ADVANTAGE (voice would be advantage to pilots' task, potential hand/eye overload) 2 = NO ADVANTAGE (voice is no advantage, equivalent to existing method of performing task) 3 = DISADVANTAGE (voice could potentially create a disadvantage to pilots performing task)
Action Criticality 1 = NO PROMPT ACTION REQUIRED (time available to repeat message) 2 = PROMPT ACTION REQUIRED (little time to verify message) 3 = IMMEDIATE ACTION REQUIRED (almost no time to verify message)
Adaptability 1 = NO PROBLEM (bidirectional data bus available, programming may or may not be required) 2 = SOME DIFFICULTY (unidirectional data bus available/in use, i.e., ARINC-429, voice system could replace or substitute for existing control head or new input/output ports could be added) 3 = MORE DIFFICULT (No data bus exists, but digital electronics are incorporated in particular aircraft system, therefore can modify particular aircraft system to add interface to voice system) 4 = VERY DIFFICULT (No data bus convenient electrical hardware exists to interface to, not worth the trouble)

Table 2.2.2 Voice Synthesis Ratings for Potential Cockpit Applications (Continued)

Potential Voice Synthesis Application	Technical Feasibility Factor	Utility/ Advantage Factor	Time/Accuracy Requirement Factor	Hardware Adaptability Factor	
Communications					
• Playback of message for Mode-S-type data transmission (N)	2	1	2	-	1(N)
• Voice record and playback of standard communications (from ATC) (E)	1	1	2	3(E)	1(N)
Navigation					
• Callout of marker beacons (E)	1	1	2	3(E)	3(N)
• MLS position information (N)	2	2	2	-	1(N)
Flight Instruments					
• Callout of airspeed and altitude (E)	1	1	2	3(E)	1(N)
Additional Aircraft Subsystems					
• Announcing alerts (E)	1	1-3	2	2-3(E)	1(N)
Flight Phase Status Monitor					
• Advanced alert system messages (N)	1	1	2	-	1(N)
• Interaction with schematics and checklists (N)	1	1	2	-	1(N)
Artificial Intelligence (AI) System					
• Interaction with AI system response (N)	2	1	1	-	1(N)

Note: (E) = Existing systems task, (N) = Next generation systems task

Table 2.2.3 Voice Recognition and Synthesis Ratings for Potential Cockpit Simulator Applications

Potential Voice Recognition and Synthesis Applications	Technical Feasibility Factor	Utility/ Advantage Factor	Time/Accuracy Requirement Factor	Hardware Adaptability Factor
Simulator Mode Control by Instructor				
• Selecting aircraft and simulator modes (REC)	1	1	1	1(E)
• Programming weather and aircraft conditions (REC)	1	1	1	1(E)
• Receiving simulator status, on request (SYN)	2	1	1	1(E)
Simulator Mode Control by Student(s)				
• Selecting aircraft and simulator modes (REC)	1	1	1	1(E)
• Programming weather and aircraft conditions (REC)	1	1	1	1(E)
• Receiving simulator status, on request (SYN)	2	1	1	1(E)
• Announcing potential hazardous flight modes or configurations (SYN)	2	1	1	1(E)

Note: Rating scheme same as used on Tables 2.1 and 2.2

The task rankings in Tables 2.2.1 and 2.2.2 merit a few comments. First, it is technically feasible to perform most of the listed tasks with a voice recognition system similar in capability to the AFTI/F-16 Phase II system, and Tables 2.2.1 and 2.2.2 reflect this. However, it would not be very practical to use such a system. Its limited vocabulary could service only one or two systems, and the isolated recognition mode would be unacceptable to pilots after a short time, if not immediately. Acceptable voice performance will be examined in task 3.

Second, the utility/advantage factor rankings indicate that the preferred use of voice, at least initially, is for programming or entering data and not for switching and mode selection. Selecting switch positions or system modes could get to be tedious work if done by voice control unless more efficient access and entry modes are developed. The pilots indicated that entering data, e.g., frequencies and way points, by voice would be an option worth considering. Because the expressed preference for voice recognition was for entering data (multiple word entries), a 90% minimum and 95% desired overall recognition accuracy was indicated as necessary for acceptable operation of these tasks. From observing pilot responses to a system with variable accuracy, it is estimated that 95% to 98% minimum accuracy may become a requirement in order to meet pilot expectations for normal operations.

Finally, the hardware adaptability ratings indicate that most of the subsystems that show promise in existing commercial aircraft use ARINC-429 interfaces. These data buses are not the most convenient to use but, depending on the subsystem, a voice system could be interfaced to many of them. It is assumed that next-generation commercial aircraft will be using high-speed bidirectional data buses and interfacing will be much easier.

2.2.2.2 Interfacing With New Aircraft Systems

A number of new aircraft systems were also considered for control or information exchange by voice recognition and/or synthesis. These systems are listed in Figure 2.2-1 as possible future aircraft subsystems. The new systems will most likely use digital computers and have a high-speed bidirectional data bus available for a voice system interface. The identification and ranking of potential tasks associated with these systems are also listed, together with existing systems, in Tables 2.2.1 and 2.2.2. The same ranking criteria was used.

Voice recognition tasks associated with the new systems generally involve mode selection or programming. Mode selection with voice may be advantageous at some times, but generally it has no direct advantage over standard manual switching schemes. Workload improvements are possible if accuracy is sufficiently high. Also, voice systems can be designed to have advantages over complex control functions in automated or semiautomated systems. Other possibilities include conditions involving extensive visual attention when mode switching might be desired, such as changing symbol modes on a head-up display (HUD) during final approach - to refine the display or to change to a go-around mode. Systems programming by voice recognition methods is promising, especially if connected or continuous speech recognition is available.

Voice synthesis tasks have been concluded to be useful with proposed alerting and interactive systems. It has been recommended that all voice alerts be carefully implemented in accordance with alerting system guidelines recently developed by the FAA and major commercial aircraft manufacturers (ref. 6). Questions to be resolved include use of voice response with the interactive systems in such a way that it does not interfere during alert situations. All those contributing to the task rankings specified this as a necessary requirement for pilot acceptance and safety.

2.2.2.3 Potential for Use in Pilot Training

The use of voice systems in training simulators shows great potential in setting up conditions and in providing immediate pilot feedback. Voice recognition could be used by the operator to select a specific aircraft or weather condition and change it directly from the simulator pilot's seat. Voice synthesis could be used for reporting aircraft, weather, or simulator status. A simulator voice system could also be used by an instructor or by certified pilots performing check flights. All this can be done with commercial voice recognition and synthesis systems available today.

Modern training simulators approach the sophistication of airplanes; also, many airports and interconnecting routes are available in the simulator's data banks. They offer the instructor a wide variety of aircraft and weather conditions to choose from; however, there are operational complexities. For example, to select the desired conditions for a "flight," an instructor must step through several "pages" on the simulator control screen and enter information here and there. Voice recognition could be used to jump directly to the "page" of interest and enter data. This could be done by the instructor at the console or in one of the pilots' seats. Similarly, voice synthesis could be used to present status on request.

For a better understanding of how voice systems could contribute to a commercial flight deck operation, a B-767 fixed based training simulator was used to provide a frame of reference for discussions. A few typical short commercial routes were flown to get a feeling for the extent current generation cockpits have been automated. It became apparent that the training simulator's operation could also benefit from voice recognition and synthesis applications.

The pilots were interested in using a voice system to program the simulator and receive status information. Table 2.2.3 reflects the positive nature of their ratings. Additionally, one of the instructor pilots suggested that the same voice system could be similarly useful to the airlines in pilot proficiency check flights. The pilots could be given simple instructions as to what aircraft, weather, and location conditions should be used for their check ride. The simulator could then be configured by interactive voice control. Besides cutting down on the number of instructors, the pilots would probably be more at ease and be able to learn more about their airplane by flying in various aircraft configurations and weather conditions.

Although an isolated voice recognition system could be used in simulators, a connected or continuous word system would be more acceptable and easier to use. Flight-quality systems would not be required and there are several commercial systems currently available that could be used. The simulators presently use high-speed bidirectional data buses, so interfacing the voice systems with the simulator would not be difficult.

Training applications could require considerably more words than flight operations. Accordingly, although phoneme-based speech synthesis is not necessary, there is potentially a large vocabulary which would be more cost effective and easier to change with phoneme-based speech synthesis as the simulator is periodically updated.

2.3 Task 3: Suitability of Voice Systems for Use in a Commercial Aircraft Operating Environment

The previous portion of the study effort was to develop a list of benefits and constraints for voice applications in commercial cockpits. With these benefits and constraints in mind, the earlier ratings will be adjusted in terms of the practicality and desirability of each potential task.

2.3.1 Environmental Constraints on Use of Voice in the Cockpit

The environment that the voice systems will have to operate in will be considered first. Electronic equipment designed for commercial aircraft will be required to meet specifications to be developed by the organization formed to support Airlines Electronic Engineer Committee needs established by Aeronautical Radio, Incorporated (ARINC). Electrical and mechanical specifications include EMI limits, operating and storage limits for temperature and pressure, power requirements, equipment case design and size, etc. The ARINC electrical and mechanical requirements should not pose any real problems to voice system manufacturers. As indicated earlier, several voice manufacturers have already developed systems that meet more stringent military qualifications for the AFTI program.

When ARINC addresses voice recognition requirements specifically, an environmental constraint to be expected is the operational minimum signal-to-noise ratio of voice to ambient noise. Voice recognition systems without speech enhancement capability need a signal-to-noise ratio of about 20 dB. Obviously, ambient noise and microphone type and location will affect the quality of the voice signal getting to the voice system.

The ambient noise that the AFTI/F-16 voice systems have to operate in (up to 115 dB) is far above that of the commercial cockpit. In new-generation jets such as the B-767 and B-757, the ambient cockpit noise level in the typical speech range (0.5k to 4k Hertz) ranges from 60 to 76 dB (Boeing B-767, B-757 sound level documents 1982 and 1984). With the relatively low ambient noise, most recognition systems can operate with a high degree of accuracy, especially if a directional microphone is used. Preliminary indications are that the AFTI systems operate without much degradation for noise levels up to 90 or 95 dB.

In areas where noise interferes with voice recognition, speech enhancement or noise canceling techniques such as described earlier would improve the reliability of a recognition system, especially during abnormal or emergency situations. Additionally, directional microphones will screen out other crewmember's voices as well as the ambient noise.

Conservatively it should be assumed that most, if not all, flight-certified recognition systems for the next five years will require a boom microphone (or oxygen mask microphone) for high quality operation. For current and near-term recognition systems, two sets of voice patterns would have to be stored for each pilot, one for a directional boom microphone and the other for the oxygen mask microphone. The switch that currently activates one microphone or the other could also signal a recognition system to change patterns.

Voice synthesis messages, when used in the cockpit, must be distinctive and intelligible. Synthesis introduces no new constraining factors so far as flight deck use is concerned; many of the problems were addressed in caution-warning standardization studies. The nature of the messages, type of voice, positioning of voice speaker, and amplitude all must be carefully considered (ref. 6 and ref. 7). These issues will be addressed in greater depth later in this discussion.

2.3.2 Workload Effects on Flight Deck Use of Voice Systems

Before voice systems are routinely installed in flight decks, enabling research is required to determine where voice input and/or output may offload or increase pilot workload. Much will depend on the design concept and how efficient an interface operation it provides. Present expectations are that voice systems can improve workload levels when voice input and/or output modes are optimally integrated, and when accuracy is high and substitution errors are low.

Voice recognition may not be faster for switching operations, but could be more convenient, especially for switches widely distributed throughout the flight deck. It should also be faster than keyboard operations or menu access routines, especially when using variable syntax. Utility can be further improved when the voice systems interact with expert systems that can respond to voice commands, performing commanded operations and providing feedback on progress and completion status.

During time-critical, high-workload portions of flight, selected applications of voice systems are appealing, but the real merit will be dependent upon efficiency of design for workload alleviation (both cognitive and physical). Interaction with a suboptimum voice system could actually add workload. For example, workload for voice communication and message confirmation is high during takeoff, approach, landing, and emergencies. Voice systems could be used to set up messages for data link, with the data link accomplishing the actual transfer and confirmation of correct transmittal by the receiver. The advantages are in reduced time to repeat or reconfirm messages to ensure accuracy. Other possibilities include, for example, fuel tank balancing, approach plate recovery (and display), checklist and procedures applications (including, for example, interactive voice recognition and synthesis), and reconfiguring the airplane for climb, cruise, or landing according to flight mode.

There are, however, a number of precautions to be observed. As has been suggested, voice synthesis may be less widely usable than voice recognition. Care must be exercised to avoid interference with critical messages, such as from air traffic control or from the airplane's caution-warning system. Even in caution-warning applications, it has been concluded in a Federal Aviation Agency-sponsored study that voice alerts should be automatic only in certain emergency situations (ref 6). In all other cases, voice alerting guidelines were to restrict use of voice to a pilot option; chimes are used to catch attention and visual presentation of information is used more extensively than some voice synthesis advocates might suggest. One of the key problems is that, unlike the visual scan mechanism, the auditory system is a single thread channel which strongly demands attention; there is considerable concern about information overload on the one hand, and about the possibility that excessive use will lead to familiarity and complacency on the other. Finally, there are cases where it is desirable to have ready access to any given segment of information on a task for a quick review.

Overall, there is reason to believe that voice systems can be of considerable assistance in workload alleviation but may gain little if used in simple switching, display, or menu access modes. The benefits appear in integrated applications that use recognition and feedback in a highly integrated format. Further work will be necessary to identify high payoff conditions of use for voice systems in flight deck applications. Specific concepts should be defined and appraised in a task context with examination of peak workload conditions and whether use of voice is beneficial or restrictive.

2.3.3 Benefits and Constraints of Application

A number of generalized tradeoffs for voice systems applications and constraints have been identified and are presented in Tables 2.3.1 and 2.3.2. These relate to the number of potential voice tasks which have been identified that could be performed in commercial aircraft and simulator cockpits (tables 2.2.1 and 2.2.2). The ratings assigned each task earlier herein would be helpful in selecting good applications and culling out poor ones. Some of the tasks are promising. Others are much less practical and not likely to be implemented in the near future. It remained to rate the tasks for applicability.

Ratings of benefits and constraints of application are presented in Tables 2.3.3, 2.3.4, and 2.3.5. The grouping code (e.g., 1111) is based on the net application ratings of Tables 2.2.1 through 2.2.3. For present purposes, the grouped code is correlated with a "payoff" weighting factor as outlined below. The scheme is similar to the one developed by Feuge and Geer (ref. 12) with the exception that the unassessed variable factor they use was replaced by the hard ware adaptability factor. Again, it should be noted that the technical feasibility factor that was used in Tables 2.2.1 through 2.2.3 is based on existing flight-quality voice systems and near-term expectations and not commercial-grade systems, nor are any assumptions made regarding adaptability of commercial-grade systems.

Payoff Factor = Weight	/Technical /Feasibility /Factor	/Utility and /Advantage /Factor	/Time and Accuracy /Requirement /Factor	/Hardware /Adaptability /Factor

Weighted Payoff	Net Application Ratings
1 = High Payoff	(1111, 1112)
2 = Some Payoff	(1122, 1123, 2111, 2122)
3 = Uncertain Payoff	(1211, 1212, 1213, 1222, 2212, 2213)
4 = Low Payoff	(1231, 2222, 2232, 2233)
5 = No Payoff	(X3XX)

The resulting weighted payoffs for each subsystem task are listed in the left column of Tables 2.3.3, 2.3.4, and 2.3.5. Both existing and next-generation possibilities existed for the numerous candidate applications. The choice here was to avoid dual ratings for existing versus next-generation possibilities. As before, in most cases listed in these tables, the hardware adaptability factor for existing hardware was used to compute the payoff factor. Only when there was no existing system was the next-generation factor used, e.g., Mode-S communication.

The results show, as would be expected, that similar task functions (e.g., switching) received similar payoff codes from one subsystem to the next. Top rankings for voice recognition went to programming applications. Ratings reflect less certainty regarding the use of voice for switching and mode-selecting tasks in the cockpit. Tasks that required smooth and gradual or continuous control, e.g., setting trim, were rated very low.

Voice synthesis shows a potential for a positive payoff in most tasks. Three factors contribute to this. First, good-quality voice synthesis systems (digitized and compressed) are now available and in use for aircraft applications. Second, flight-quality phoneme systems with large vocabularies and good-quality voice will be available soon. Last, in all but a few alerts, voice synthesis can be under pilot control, so presentation of information can be at the pilots' request and the voice synthesis interruption problems mentioned earlier as concerns can be avoided.

Not considered in the above-mentioned payoff factors are the practicality and desirability. To correct this the payoff factors were weighted to reflect the perceived practicality/desirability of voice in each task. The desirability of having a voice system assist in a task is quite subjective. Design features to enhance ability and efficiency of the voice system in performing a task are taken into account. Ease of use or user friendly features of the voice system also are important. All the voice tasks, with the exception of a few alerts, have yet to be tried. Therefore, ability, efficiency, and friendliness were estimated on the basis of expected capabilities.

If a proposed voice task is expected to be practical and desirable, then its payoff factor weight was reduced by one. No weight was given to those tasks that had nearly equivalent benefits and constraints or that drew little interest from the review team. Tasks that appeared to be impractical or undesirable had their payoff factor increased by one.

As shown in Tables 2.3.3, 2.3.4, and 2.3.5, discrimination of task utility has evolved. The revised task ratings are now spread out. Those tasks with the highest perceived potential have ratings of 0 and 1. These ratings will be used later herein to develop a benefits hierarchy of cockpit voice tasks.

Table 2.3.1 Voice Recognition Benefits and Constraints

Benefits	Constraints
• In tasks where pilot attention must be diverted during a critical part of flight, voice can provide the ability to reduce workload • Could speed data entry, especially when hands and/or eyes are busy	• Acceptability and usefulness depends on high overall accuracy and errors being primarily rejection, not substitution
• Several existing systems could be interfaced and controlled as indicated in Tables 2.2.1-2.2.3	• ARINC-429 bus is unidirectional and not convenient to interface • Interfacing with numerous systems may be difficult on existing aircraft • Selecting system modes with voice is also of uncertain advantage
• Radios could be tuned by saying locale identification	• Tuning radios by entering frequency may or may not be any improvement over manual method
• An active vocabulary of 50 words with paging can handle most aircraft systems and can transfer voice reference data files with host system	• Flight-qualified systems with this capability will not be ready for 1 to 3 years • Systems will be speaker dependent require two or more training passes per word • Vocabularies must be carefully selected to avoid similar sounding words • Flexible syntax necessary to enhance user acceptance • Each pilot will need own system and reference data
• Flight-qualified systems available now from at least four vendors	• Existing flight-qualified systems are speaker dependent, isolated mode, limited vocabulary, and expensive. Microphone must be located close to mouth, boom microphone or oxygen mask
• Commercial quality connected and continuous systems available now. Some may be flight-quality, pending evaluation	• Training for connected and continuous modes is tedious • Some report errors when receiving nontrained words
• It is technically feasible to produce flight-qualified connected mode voice recognition systems now	• Sophisticated host required for real-time response, flexible syntax, and multisystem control

Table 2.3.2 Voice Synthesis Benefits and Constraints

Benefits	Constraints
• Useful for some aircraft alerts • Potential for recording and playback of messages • Potentially useful in interacting with pilots for checklists, etc.	• Voice can add to confusion in emergency situations and decrease safety
• High-grade commercial and flight-quality digitized and compressed synthesis systems now available. Boards or chipsets	• Digitized and compressed voice systems have limited active vocabulary, each word has to be recorded, and sound of word is dependent on how it was recorded
• Good-quality commercial phoneme systems with more than 10K word vocabulary and speech options now available • Technically feasible to make flight quality phoneme systems • Speech options on some phoneme systems permit varying voice types and rates	• No flight-quality phoneme systems available now • Phoneme speech quality not yet equal to digitized and compressed methods

Table 2.3.3 Revised Ratings for Potential Cockpit Applications of Voice Recognition

Potential Voice Recognition Task	Summary of Task 2 Ratings	Desirability Weighting	Revised Task Rating
Communications			
• Switching and selecting modes (E) (IWR)	3	+0	=3
• Volume control (E) (IWR)	5	+1	=6
• Entering frequencies (E) (IWR)	2	+0	=2
• Radio tuning by location ID (E) (IWR)	2	−1	=1
• Selecting and preparing messages for Mode-S-type data transmissions (N) (CWR or IWR)	2 or 3[1]	−1 or 0	=1 to 3
Navigation			
• Switching and selecting modes (E) (IWR)	3	0	=3
• Entering frequencies (E) (IWR)	2	+0	=2
• Radio tuning by location ID (N) (IWR)	2	−1	=1
• Programming CDU [IRS, NAV and performance management systems] (E) (CWR and IWR)	2 or 3[1]	−1 or 0	=1 to 3
• Programming microwave landing system (N) (CWR or IWR)	2 or 3[1]	−1 or 0	=1 to 3
Flight Controls			
• Primary attitude controls (E) (CWR)	5	+1	=6
• Selecting positions for flaps, speed brakes, and trim (E) (IWR)	5	+0	=5
• Selecting autopilot and fuel management systems modes (E) (IWR)	3	+0	=3
• Entering data to autopilot and thrust management computer (E) (IWR)	2	+0	=2
• Selecting modes for autopilot and advanced fuel management system (N) (IWR)	3	+0	=3
• Programming 4-D navigation system (N) (CWR)	2	−1	=1

NOTES: (1) If a connected word recognition system is used, the task 2 summary rating is better (a lower number) than if isolated word recognition is used.
(2) The ratings from Table 2.2.1 vary depending on which subsystem voice recognition would be used.
(3) (E)=existing systems task, and (N)=next-generation task
(4) (IWR)=isolated word recognition, (CWR)=connected word recognition

Table 2.3.3 Revised Ratings for Potential Cockpit Applications of Voice Recognition (Continued)

Potential Voice Recognition Task	Summary of Task 2 Ratings	Desirability Weighting	Revised Task Rating
Flight Instruments			
• Selecting speed and height bugs (E) (IWR)	2	+0	=2
• Entering barometric pressure (E) (IWR)	2	+0	=2
• Entering modes for EADI, EHSI, EICAS, and HUD (E/N) (IWR)	3	+0	=3
Additional Aircraft Subsystems [hydraulics, electrical, pneumatics, fuel, air conditioning, engines, APU, anti-ice, rain protection, fire protection, landing gear, crew alerting]			
• Selecting positions and modes (E) (IWR)	3 to 5^2	+0	=3 to 5
• Integrated systems management (N) (CWR or IWR)	2 or 3^1	–1 or 0	=1 to 3
Flight Status Monitor			
• Interaction with schematics and checklists (N) (CWR or IWR)	3 or 4^1	–1	=2 to 3
• Request status (N) (CWR or IWR)	2 or 3^1	–1 or 0	=1 to 3
Programmable Multipurpose Keyboard			
• Paging (N) (CWR or IWR)	2 or 3^1	–1 or 0	=1 to 3
• Entering data and programming aircraft subsystems like CDU (N) (CWR or IWR)	2 or 3^1	–1 or 0	=1 to 3
Multipurpose Displays			
• Paging and format request (N) (CWR or IWR)	2 or 3	–1 or 0	=1 to 3
• Request and step-through operations checklists (N) (CWR or IWR)	2 or 3	–1 or 0	=1 to 3
Artificial Intelligence (AI) System			
• Interaction with AI system's recognition/understanding (N) (CWR)	2	–1	=1

NOTES: (1) If a connected word recognition system is used, the task 2 summary rating is better (a lower number) than if isolated word recognition is used.

(2) The ratings from Table 2.2.1 vary depending on which subsystem voice recognition would be used.

(3) (E)=existing systems task, and (N)=next-generation task

(4) (IWR)=isolated word recognition, (CWR)=connected word recognition

Table 2.3.4 Revised Ratings for Potential Cockpit Applications of Voice Synthesis

Potential Voice Synthesis Task	Summary of Task 2 Ratings	Desirability Weighting	Revised Task Rating
Communications			
• Voice record and playback of standard communications from air traffic control (E) (DIG)	2	-1	=1
• Playback of messages for Mode-S-type data transmission (N) (PHO)	2	-1	=1
Navigation			
• Callout of marker beacons (E) (DIG)	2	+0	=2
• Callout of position information from microwave landing system (N) (PHO)	4	+0	=4
Flight Instruments			
• Callout of airspeed and altitude (E) (DIG)	2	-1	=1
Additional Aircraft Subsystems			
• Announcing alerts (E) (DIG)	2	-1	=1
Flight Status Monitor			
• Advanced alerting system message (N) (PHO)	2	-1	=1
• Interaction with schematics and checklists (N) (PHO)	2	+0	=2
Artificial Intelligence (AI)			
• Interaction with AI system, response (N) (PHO)	2	-1	=1

Notes: (1) (E)=existing systems task, (N)=next generation task
(2) (DIG)=digitally compressed synthesis, (PHO)=phoneme synthesis

Table 2.3.5 Revised Ratings for Potential Simulator Applications of Voice

Potential Voice Recognition and Synthesis Tasks	Summary of Task 2 Ratings	Desirability Weighting	Revised Task Rating
Simulator Mode Control by Instructor:			
• Selecting aircraft and simulator modes (REC) (E) (IWR)	1	+0	=1
• Programming weather and aircraft conditions (REC) (E) (IWR)	1	+0	=1
• Receiving simulator status on request (SYN) (E) (PHO)	2	+0	=2
Simulator Mode Control by Student(s):			
• Selecting aircraft and simulator modes (REC) (E) (IWR)	1	+0	=1
• Programming weather and aircraft conditions (REC) (E) (IWR)	1	+0	=1
• Receiving simulator status on request (SYN) (E) (PHO)	2	+0	=2
• Announcing potential hazardous flight modes or configurations (SYN) (N) (PHO)	2	+0	=2

Notes: (1) (E)=existing systems task, (N)=next generation task
(2) (REC)=recognition mode, (SYN)=synthesis mode
(3) (IWR)=isolated word recognition, (CWR)=connected word recognition
(4) (DIG)=digitally compressed synthesis, (PHO)=phoneme synthesis

2.3.4 Flight Deck Voice Recognition Performance Considerations

Primary performance considerations to achieve ultimate acceptance will be accuracy and error types. However, besides accuracy and error requirements, there are a number of additional factors that must be considered. These factors include ambient noise, microphone type, vocabulary make-up and size, and syntaxing controls.

The recognition tasks that show the most promise as a result of the ratings in Tables 2.3.3 and 2.3.5 are estimated to require, eventually, 95% to 98% or better isolated mode accuracies in repeatability in order to achieve eventual acceptance by line pilots. Although the acceptable levels for other error types have not been mentioned, they are as important as the repeatability accuracy levels. False rejection is preferable to substitution or false acceptance errors, but acceptable ratios of these errors is dependent on the task and its urgency.

Ambient noise level high enough to mask a operator's voice will make a voice system useless. This should not be a problem in the low noise environment (60-76 dB in the 757 and 767) of late model commercial aircraft cockpits, especially if condenser-type boom microphones are used. If hand-held microphones are used, then voice patterns as seen by a recognizer will differ if the microphone is not held in exactly the same position each time. Boom microphones are commonly used in commercial aircraft; therefore, requiring them for use with recognition systems should not be a problem.

Vocabulary selection and use can make a big difference in recognition performance with existing and near-term recognition systems. When building vocabularies, designers must be careful to exclude acoustically similar words. If two or more similar words are necessary, then no two should be active at any one time. Also, the active vocabulary should be kept small. The fewer words a system has to choose from, the faster its response and higher its chances of selecting the correct one.

The preliminary aircraft subsystem vocabulary sets (outlined in sec. 2.2.1) are expected to be less than 50 words for each subsystem any one flight. The one exception would be some of the extended word messages being considered for data link. Most likely a task would need only 10 to 15 words active at any one time. However, pilots may not be inclined to say words in exactly the same sequence. Therefore, a system with syntaxing capability should increase the overall accuracy and speed of response.

When a task's representative vocabulary and syntaxing scheme have been developed, candidate recognition systems should be tested by pilots in an ambient noise and operations environment typical of the intended flight deck. Rejection thresholds, substitution, and false acceptance error rates and, in turn, the rejection error rate and overall accuracy should be established as should effects of errors and efficient workarounds to avoid or correct errors.

Recognition accuracy and speed could be improved, if required, by using a syntax controller to manage syntaxing schemes, anticipate necessary vocabulary, and examine word strings to assure that syntactically correct combinations have been received from (recognized by) the recognition system. Texas Instruments has designed a program that incorporates such a controller for the Rome Air Development Center (Air Force) that operates with the TI-PC voice system and can handle a total vocabulary of around 300 words.

In another approach to improve accuracy, some recognition systems provide a second choice word with each word that has been selected. In an error situation the second choice word is examined, and used if it fits the syntax requirements. Such a syntaxing program was designed and built for the Navy to improve airborne voice recognition and synthesis systems' performance (ref. 33).

Controlling voice recognition/synthesis systems is a prime area for applying artificial intelligence (AI) systems. AI will almost assuredly be integral within voice systems by 1990, if not before.

2.3.5 Flight Deck Voice Synthesis Performance Considerations

Performance of voice synthesis devices is measured in terms of intelligibility, distinctiveness from competing voice and noise sources, and ability to replicate human voice qualities. The capability of voice synthesis systems to accommodate large vocabularies should also be considered after viewing the number of potential voice synthesis tasks identified in Tables 2.3.4 and 2.3.5. The two methods of voice synthesis (digitally compressed and phoneme based) considered to have potential in aircraft cockpits possess both positive and negative characteristics. These characteristics must be weighed and then compared with the task requirements.

Development of the voice model should follow the precedent set by caution-warning program studies: Berson, et al (1981) recommended that the desired characteristics of aircraft voice alerts should be chosen through empirical testing in a representative ambient noise environment to ensure intelligibility and distinctiveness. Two standard tests for intelligibility were suggested as relevant for the cockpit environment by Berson, et al, the Modified Rhyme test and the Harvard Phonetically Balanced Word test. One or both of these tests can be used to compare voice models (digitally compressed and/or phoneme type) in typical ambient cockpit noise environment. Once final candidate systems are selected, laboratory or simulator tests should be conducted with a set of voice messages from the target cockpit's tasks. Representative normal and abnormal ambient noise should be present (alternately) throughout the test.

Voice synthesis systems that use digitally compressed voice can be very intelligible, depending on the quality of the digitized voice model. A good model requires a controlled recording environment, professional recording equipment, and a trained speaker who has an intelligible voice. However, distinctiveness of a voice model in the intended environment is just as important as intelligibility. If the voice is not distinctive in the cockpit voice and noise environment, then it may be confused with other voice communications or missed altogether.

The digitally compressed voice synthesis method is currently used in some military and commercial aircraft cockpits for alert messages. This method was selected over phoneme type voice synthesis because it offered better intelligibility, distinctiveness, and overall voice quality. Recent advances in phoneme voice synthesis have greatly reduced this advantage.

One area where phoneme synthesis has a particular advantage is in potential vocabulary size. Digitally compressed techniques typically require 2400 to 9600 bits of memory per second of speech. Thus, 30 seconds of speech would require a minimum of 9K bytes of memory. A large vocabulary will quickly run up large memory requirements. Additionally, digitized systems enunciations are restricted by the way a voice message was reduced. Alternatively, phoneme systems require just a few bytes to define each word, resulting in a much smaller word/bytes-to-memory ratio than digitally compressed techniques. Another advantage to the phoneme-based system is the ability to vary inflection, speed, and pitch of speech. For example, the phoneme based DEC-talk system by Digital Equipment, Inc. uses a single printed circuit board to produce vocabulary of over 10K predefined words. It also allows a host system to control voice types (male/female/child), speaking rate, pronunciation, and intonation.

The technical feasibility factor ratings in Tables 2.3.4 and 2.3.5 indicate that digitally compressed voice recordings would be adequate in most individual tasks. However, if one voice synthesis system is to handle more than one task it becomes less practical to use digitized voice playback, because of the cumulative memory required for the combined vocabularies. A single phoneme type system could support all the cockpit tasks, although a separate one for the alerting system may be necessary.

2.3.6 Pilot-Based Implementation Guidelines

After voice systems' benefits and constraints have been considered, specific applications have been identified, and performance requirements have been defined, the next major step comes with making the transition from the laboratory to the flight deck (simulator or aircraft). Pilots' acceptance or rejection of the voice applications will depend on several criteria that must be examined at this point.

In addition to a recognition system's actual performance in the cockpit, as noted in Section 2.3.4, the other areas to be taken into account include influence of pilots' knowledge and proficiency on performance, hardware reliability, corrective actions available, clarity of feedback, guidance through the task operations without manuals, flexibility of syntaxing, transparency of menu paging, and failure mode provisions. Each of these areas will be expanded upon below, followed by a similar discussion of voice synthesis criteria.

- Recognition performance should be evaluated by pilots in the target cockpit environment with a representative message set

Minimum acceptable accuracy and error rates were proposed in Section 2.3.4. A designer should conduct a test with pilots who will use the target cockpit so that acceptable accuracies and error rates can be defined.

- Errors by the system or pilot must be easy to correct and not require repeating the entire command string

Mistakes will occur. They may be from substitution or insertion errors, false rejections of correct words, or the pilot initiating a task that he or she wants to terminate. The simplest way to handle errors is to have the pilot repeat the entire command. This is very time consuming and will probably evoke a negative reaction from the pilots whenever errors occur. A syntax monitor and controller, as mentioned in Section 2.3.4, could detect errors and permit the pilot to correct just the invalid portion of the message. If the system has a preentry display (visual or auditory), the pilot could confirm commands before actions are taken while mistakes are easily caught and corrected. Single-line displays or portions of larger displays could be used (e.g., CDU scratch pad or engine instrument displays). Preentry visual displays have been used in simulator evaluations of voice systems at Boeing for a number of years.

- The recognition system must clearly inform the pilot what action, if any, has been taken

As noted above, a preentry display would allow the pilot to verify that correct commands have been received before approval is given to execute them. When the system detects errors it must inform the pilot in a clear, concise manner and prompt a corrective action.

- Guidance should be provided to the pilot so that operations manuals are not necessary

On tasks where the pilot may enter and program at different levels of an information/data tree, some visual display must be provided to inform what actions are being taken and what system is being operated on. For example, if the navigation system is to be modified with the voice system, then the CDU display should present the navigation page and associated data that is being operated on. If multifunction switches are used, their displays should show information corresponding to the displayed page. Some guidance may also be given via a preentry display.

- Direction and interrogation messages should be consistent with the vocabulary commonly used by pilots

- Variable syntax should be incorporated if possible

Variable syntax messages are possible with systems that use a recognition controller (sec. 2.3.4). Pilots are more likely to accept and use a system that is flexible in the way messages or commands are spoken. This flexibility will in effect raise the overall operational accuracy and make the system more useful in high-stress situations when exact word arrangements of commands may be forgotten.

- System performance should remain constant for all phases of flight

As with all other systems on the flight deck, a recognition system must perform uniformly in normal, abnormal, and emergency flight conditions. The environmental conditions that an airborne recognition system will have to operate in are noted in Section 2.3.1.

- Failure mode provisions must be provided and not degrade aircraft control

If a recognition system experiences a failure or abnormality, the pilot must be made aware immediately. The failure must not lock up or hinder any other system in the cockpit. Also, a parallel method of executing a task must be available. For example, if some coordinates are being entered into the navigation system with voice, the CDU is configured accordingly. Thus, if the voice system fails, the pilot may continue entering the coordinates via the CDU keyboard.

- Training on the recognition vocabulary should be able to be conducted in or out of the cockpit

The pilots should have the option of training with the recognition vocabulary in the cockpit or at a ground-based training facility. The AFTI/F-16 voice command systems have this option.

- Training sessions should be self-prompting and inform the pilot if any words or phrases should be retrained

Training will involve speaking each vocabulary word two or more times. The recognition training software routine should prompt the pilot with the word to be spoken, e.g., "PLEASE SAY *WAY POINT*." If the recognizer has difficulty identifying any trained words it should request that the pilot retrain on those words.

- The pilot should be able to update/retrain any word or phrase in or out of the cockpit

The pilot should have the option to update/retrain any vocabulary word while in the cockpit or at a training facility. If a small number of words are not being recognized repeatedly, a pilot may desire to update the voice reference patterns for those words. This would be more convenient than retraining all the vocabulary words.

Voice synthesis acceptance features will include the performance requirements noted in Section 2.3.5 and whether it interferes with normal, abnormal, and emergency cockpit activities.

- Synthesis performance should be evaluated by pilots in the target cockpit environment with a representative message set

Cockpit voice synthesis performance requirements were discussed in Section 2.3.5

- Interferences with cockpit activities should be avoided

2.3.7 Comparison of Cost Factors: Voice Versus Hardware

To start out with, voice systems will not replace any hardware systems in the aircraft or simulator cockpit. At least not in the near future. As noted in Section 2.3.6, initial applications of voice systems should parallel rather than replace input-output methods for existing or new cockpit systems. The "cost savings" will be in reduced workload and increased safety.

The cost of voice systems will be twofold; the cost of the actual systems and costs involved with interfacing them to other aircraft or simulator systems. Figure 2.3.7 lists price quotes from two manufacturers for limited production of existing flight-quality voice recognition and synthesis systems. Prices for these military-quality flight systems are notably higher than those for other commercially used systems, e.g., Appendix A. A significant part of this is presumed to be related to the more stringent requirements for the military-qualified system; additionally, the systems are not being produced in large quantitites. However, while it is suspected that other systems would be suitable for the far more benign commercial aircraft environment, insufficient data are available to confidently identify suitable candidates for test in this more compatible environment. While the cost of these limited production items is relatively high, if a commercial and/or military market develops and quantity lots are ordered, the prices will likely drop. While prices are not likely to plummet, the range of costs in Appendix A gives an appreciation of what can happen in this very competitive field.

Some commercial-quality (not flight-qualified) systems have been tested or are planned to be used in test in experimental aircraft to evaluate potential for application of voice systems in performing flight deck tasks. They have been assessed as sufficiently durable for experimental evaluations of task performance on an experimental aircraft if appropriate precautions are taken; they may be of flight quality though not flight qualified. Thus, they offer a lower cost alternative for exploratory evaluation to define both task applicability and the detailed requirements for commercially flight-qualified voice recognition and synthesis systems. These systems include those produced by Interstate Electronics Corporation and Texas Instruments.

ITT Defense Communications Division, San Diego, California		
	One Unit	Three Units
ITT VCS	$50,000	$150,000
The ITT Voice Command System (VCS) is configured similarly to their AFTI system but with a 16-bit parallel interface instead of Mil. Std. 1553 and without safety of flight testing requirements.		
Texas Instruments, Inc., Equipment Group, Dallas, Texas		
	One Unit	Three Units
TI VIS	$256,000	$410,000
The TI Voice Interactive System (VIS) is configured similarly to their AFTI system but with a 16-bit parallel interface instead of Mil. Std. 1553. The modified system would undergo acceptance test in accordance with AFTI test procedures.		

Note: These prices reflect small quantities of flight certifiable equipment

Figure 2.3.7 Budgetary Price Quotes for Two Voice Recognition Systems Designed to Military Qualifications Tests, April 1984

The cost to interface voice systems will depend on the type of data bus used and the extent to which they parallel other systems. If a monodirectional bus such as the commercial ARINC-429 bus is used, each voice system may need one input and one output port for each device it is to control! Also, each controlled device has to add input and/or output ports and be reprogrammed to interact with the new ports. The cost to do this would obviously be high.

However, to evaluate a voice in an existing test aircraft, one redundent control device, e.g., EHSI or control display unit (CDU), could be replaced by a properly programmed voice system. If the voice recognition system could work in parallel with another control device , e.g., CDU, then only that device would have to be reprogrammed and interfaced to. For near-term exploratory applications this would probably be the most cost efficient. Similarly, voice synthesis system could be interfaced to an existing alerting system.

Next-generation aircraft are expected to incorporate bidirectional, high-speed parallel or serial data buses. A voice system could attach to such a bus and control numerous cockpit subsystems. Multifunction keyboards and voice systems could operate in parallel, tracking each other, and the controlled subsystem would not know or care which keyboard or voice system gave the command.

2.4 Task 4: Identification and Recommendation of Cockpit Voice Applications

Task 4 assimilates the information generated in the first three tasks into a hierarchical benefits rating of potential voice applications, a list of general purpose design guidelines, and proposals for five generic voice cockpit and simulator voice systems.

2.4.1 Benefits Hierarchy of Potential Cockpit Voice Tasks

This section culminates the rating efforts from the previous two tasks into a benefits hierarchy of potential cockpit voice tasks.

A number of voice recognition and synthesis tasks were identified in task 2 and rated on four criteria: (1) the minimum level of voice technology required to perform the task; (2) the advantage a voice system would offer pilots over existing methods; (3) the accuracy/intelligibility necessary to perform the task considering its urgency and time available to correct errors; (4) the adaptability of interfacing the voice system to the aircraft subsystem associated with that task (tables 2.2.1 through 2.2.3).

This rating scheme helped identify which voice tasks may have promise and those that do not; however, a number of other factors that could affect the success of a voice system at each task were not rated. To take these into account, each task's four ratings were summarized into a payoff factor (high, some, uncertain, low, and no payoff). The merit of each task was evaluated for overall desirability and its payoff code was weighted accordingly (tables 2.3.1 through 2.3.3). Desirability was based on environmental constraints, pilot workload considerations, and the benefits and constraints listed in Tables 2.3.4 and 2.3.5. It is from the revised payoff ratings listed in Tables 2.3.1 through 2.3.3 that a benefits hierarchy was drawn.

Tables 2.4.1 through 2.4.3 list the potential voice tasks and their benefits hierarchy ratings. The rating scheme follows the one for payoff factors in Section 2.3.3.

CODE = 1, FOR HIGHLY BENEFICIAL VOICE TASK
= 2, FOR SOME BENEFIT TO VOICE TASK
= 3, FOR UNCERTAIN IF BENEFICIAL VOICE TASK
= 4, FOR LOW BENEFIT TO VOICE TASK
= 5, FOR NO BENEFIT TO VOICE TASK

The code numbers are essentially the same as the revised ratings listed in Tables 2.3.1 through 2.3.3 except for a few tasks which would have resulted in revised ratings of 0 and 6; these were assigned a benefits codes of 1 and 5 respectively. Also, tasks that had multiple ratings, e.g., programming CDU, had the best (lowest) rating passed on to the benefits hierarchy tables.

Table 2.4.1 proposes that voice recognition's best use in commercial aircraft cockpits is to program and interrogate complex systems. Programming would include identifying a specific system and commanding a parameter or plan be changed, e.g., "update NAV system - select ILS approach to Seattle via runway 34." Programming, as used here, would also include setting up several common systems simultaneously with one command, e.g., "set comm radios to Portland approach control." Interrogating the status of a system would provide the pilot information about a system immediately without having to page through several menus on a CDU. It is assumed that such a recognition system will operate in parallel with another data entry device (e.g., multifunction keyboard or touch screen display) and have feedback from one or more sources (e.g., preentry or multifunction displays). Also, recognition systems associated with programming tasks must have high accuracy (estimate ≥98% overall) and operate in connected or continuous recognition modes. Otherwise, pilots will have to repeat commands or correct errors too often. If this happens, they will just turn the system off!

Table 2.4.1 Benefits Hierarchy of Potential Cockpit Applications of Voice Recognition

Potential Voice Recognition Task	Benefits Ratings
Communications	
• Switching and selecting modes (E) (IWR)	3
• Volume control (E) (IWR)	5
• Entering frequencies (E)(IWR)	2
• Radio tuning by location ID (E) (IWR)	1
• Selecting and preparing messages for Mode-S-type data transmissions (N) (CWR)	1
Navigation	
• Switching and selecting modes (E) (IWR)	3
• Entering frequencies (E) (IWR)	2
• Radio tuning by location ID (N) (IWR)	1
• Programming CDU [IRS, NAV and performance management systems] (E) (CWR)	1
• Programming microwave landing system (N) (CWR)	1
Flight Controls	
• Primary attitude controls (E) (CWR)	5
• Selecting positions for flaps, speed brake, and trim (E) (IWR)	5
• Select autopilot and fuel management systems modes (E) (IWR)	3
• Entering data to autopilot and thrust management computer (E) (IWR)	2
• Selecting modes for autopilot and advanced fuel management system (N) (IWR)	3
• Programming 4D navigation system (N) (CWR)	1

Lowest values=best ratings

Table 2.4.1 Benefits Hierarchy of Potential Cockpit Applications of Voice Recognition (Continued)

Potential Voice Recognition Task	Benefits Ratings
Flight Instruments	
• Selecting speed and height bugs (E) (IWR)	2
• Entering barometric pressure (E) (IWR)	2
• Selecting modes for EADI, EHSI, EICAS, and HUD (E/N) (IWR)	3
Additional Aircraft Subsystems [Hydraulics, electrical, pneumatics, fuel, air conditioning, engines, APU, anti-ice, rain protection, fire protection, landing gear, crew alerting]	
• Selecting positions and modes (E) (IWR)	3
• Integrated systems management (N) (CWR)	1
Flight Status Monitor	
• Interaction with schematics and checklists (N) (CWR)	2
• Request status (N) (CWR)	1
Programmable Multipurpose Keyboard	
• Paging (N) (CWR)	1
• Entering data and programming aircraft subsystems like CDU (N) (CWR)	1

Lowest values=best ratings

Table 2.4.2 Benefits Hierarchy of Potential Cockpit Applications of Voice Synthesis

Potential Voice Synthesis Task	Benefits Ratings
Communications	
• Voice record and playback of standard communications from air traffic control (E) (DIG)	1
• Playback of messages of Mode-S-type data link transmission (N) (PHO)	1
Navigation	
• Callout of marker beacons (E) (DIG)	2
• Callout of position information from microwave landing system (N) (PHO)	4
Flight Instruments	
• Callout of airspeed and altitude (E) (DIG)	1
Additional Aircraft Subsystems	
• Announcing alerts (E) (DIG)	1
Flight Status Monitor	
• Advanced alerting system messages (N) (PHO)	1
• Interaction with schematics and checklists (N) (PHO)	2
Artificial Intelligence (AI)	
• Interaction with AI system response (N) (PHO)	1

Lowest values=best ratings

NOTES: (1) (E)=existing systems task, (N)=next generation task
(2) (DIG)=digitally compressed synthesis, (PHO)=phoneme synthesis

Table 2.4.3 Benefits Hierarchy of Potential Simulator Applications of Voice

Potential Voice Recognition and Synthesis Tasks	Benefits Ratings
Simulator Mode Control by Instructor	
• Selecting aircraft and simulator modes (REC) (E) (IWR)	1
• Programming weather and aircraft conditions (REC) (E) (IWR)	1
• Receiving simulator status on request (SYN) (E) (PHO)	2
Simulator Mode Control by Student(s)	
• Selecting aircraft and simulator modes (REC) (E) (IWR)	1
• Programming weather and aircraft conditions (REC) (E) (IWR)	1
• Receiving simulator status on request (SYN) (E) (PHO)	2
• Announcing potential hazardous flight modes or configurations (SYN) (N) (PHO)	2

Lowest values=best ratings

NOTES: (1) (E)=existing systems task, (N)=next generation task
(2) (IWR)=isolated word recognition, (CWR)=connected word recognition
(3) (DIG)=digitally compressed synthesis, (PHO)=phoneme synthesis

The next most likely application proposed for a cockpit voice recognition system is data entry. This function is a subset of programming. For example, a command could be given to set one radio's frequency to 123.45 or the barometric pressure reference to 29.87. In some cases voice would be useful, but often it would be just as fast or faster to set the digits manually.

Using voice commands to select subsystem modes was viewed as having uncertain value. The number of subsystems and corresponding modes, along with the varying importance of each, makes this application of voice recognition less desirable than existing manual methods. New switching methods such as multifunction switches or touch-panel displays, may be better options for improved mode selection.

Cockpit voice recognition applications that have the least benefits are those that involve critical functions and/or continuous, versus discrete inputs. For example, it is more practical and safer to adjust trim settings by hand than saying "up up up up down ... etc." Unless an alternative approach is developed, voice actuation is unlikely.

Table 2.4.2 indicates that a number of voice synthesis applications would be beneficial. It is important to note that except for a few alert situations voice response in the cockpit should be at the pilots' request. It also must be intelligible and distinctive from other cockpit voices and noise.

Table 2.4.3 proposes that voice recognition and synthesis systems would be highly beneficial for a number of applications in controlling and monitoring cockpit simulators. The advantage of voice control and monitoring in training simulators was described in Section 2.3. An important difference between adapting systems to simulators and aircraft is that commercial voice equipment can be used in simulators. Also, the conventional high-speed data buses used in simulators will be advantageous.

Research and test aircraft and simulators may also benefit from commercial voice systems. Commercial voice systems are being used in government and commercial research simulators today for evaluation of tasks similar to those noted in Tables 2.4.1 and 2.4.2. A few commercial voice systems have been and continue to be used to evaluate voice applications in government aircraft (NASA, FAA, and military).

2.4.2 General-Purpose Design Guidelines/Specifications

This section discusses some general-purpose design guidelines for implementing voice recognition and synthesis systems in aircraft and simulator cockpits and simulator control stations. These are equipment-based design criteria that a designer must consider and define along with the pilot-based guidelines (sec. 2.3.6) before choosing a voice recognition and/or synthesis system for a particular cockpit.

General Design Guidelines for Cockpit Voice Recognition Systems

- Acceptable minimum accuracy and maximum error performance levels should be determined

The pilot performance evaluation of recognition systems that was recommended in Sections 2.3.4 and 2.3.6 will yield a minimum acceptable overall recognition accuracy. Allowable percentages of the different error types should also be estimated. Recognition systems, whether isolated, connected, or continuous, should meet these specifications.

- The operational ambient noise environment should be defined or estimated for normal, abnormal, and emergency flight conditions

A recognition system must operate to at least the minimum acceptable performance levels in all modes of flight. This includes equivalent operation with boom microphone and oxygen mask. Ambient cockpit noise levels during normal commercial flights should not hamper the operation of good quality recognition systems. However, ambient noise levels during abnormal or emergency situations may become excessive (>95 dB). Therefore, noise canceling techniques should be employed so that recognition system performance is not degraded when it can be of most use.

- Direction and interrogation messages associated with each task must be defined. From this a working vocabulary can be determined

Once the cockpit systems to interact with a recognition system have been identified, a list of desired tasks should be defined. From the task functions a list of direction and interrogation messages can be formulated. The working vocabulary can then be determined. As noted in the pilot-based guidelines, the vocabulary usage and message construction should be consistent with those commonly used by pilots for the particular tasks.

- Message syntax should be noted and variable syntax combinations considered

Most messages will have two or more possible syntax combinations. If a message is likely to be stated more than one way by the pilots in normal, abnormal, or emergency situations, the system should accept the most likely combinations.

- The vocabulary/messages should be divided into minimum subsets necessary for specific tasks or group of tasks so that maximum active vocabulary sets can be defined

- Necessary response time should be appraised

- If a recognition system controller is used, then first- and second-choice words and associated ratings must be available from the recognition system

- Pilot recognition vocabulary training provisions must be defined. This should include the ability to update individual words or phrases

- Interface hardware requirements must be defined

- Interface software requirements must be defined

- Environmental conditions (besides noise) must be defined

Commercial aircraft equipment must meet environmental and structural specifications as defined by the Federal Aviation Administration.

- Storage requirements for voice records must be considered and planned

The type of recognition method employed (isolated, connected, or continuous) will depend on vocabulary size, operational environment, performance requirements, flight-quality equipment available, utility, and pilot preference. Section 2.4.1 proposes a number of potential recognition tasks and the minimum recognition method advisable for each.

General Design Guidelines for Cockpit Voice Synthesis Systems

- The operational ambient voice and noise environment should be defined or estimated for normal, abnormal, and emergency flight conditions

Characterizing the voice and noise environment must be done before tests on intelligibility or distinctiveness can be conducted.

- Acceptable minimum performance levels for intelligibility, voice quality, and distinctiveness should be determined

Voice synthesis performance criteria were discussed in Sections 2.3.5 and 2.3.6. The results of criterion tests, along with pilot evaluation, will define what performance levels are necessary for the target cockpit. These evaluations should be conducted in a voice and noise environment noted in the previous guideline.

- Voice synthesis messages associated with each task must be defined. From this a working vocabulary can be determined

Once the cockpit systems to interact with a synthesis system have been identified, a list of desired tasks should be defined. From the task functions a list of messages can be formulated. The working vocabulary can then be determined. As noted in the pilot-based guidelines, the vocabulary and message construction should be consistent with those commonly used by pilots for the particular tasks.

- Presentation of voice synthesis messages to pilots must be planned

- The method for the pilots to request voice synthesis information must be defined

- Interface hardware requirements must be defined

- Interface software requirements must be defined
- Environmental conditions (besides noise) must be defined
- Storage requirements for voice records must be considered and planned

The type of synthesis method employed (digitally condensed or phoneme) will depend on vocabulary size, operational environment, performance requirements, flight-quality equipment available, utility, and pilot preference. Section 2.4.1 proposes a number of potential synthesis tasks and the methods recommended for each.

A summary of both the pilot-based and general design guidelines is listed in Tables 2.4.4 and 2.4.5.

Table 2.4.4 Summary of Design Guidelines for Cockpit Voice Recognition Systems

Pilot Based Design Guidelines for Voice Recognition Systems (sec. 2.3.6)

- Recognition performance should be evaluated by pilots in the target cockpit environment with a representative message set
- Errors by the system or pilot must be easy to correct and not require repeating entire command strings
- A recognition system must clearly inform the pilot what action, if any, has been taken
- Guidance should be provided to the pilot so that operations manuals are not necessary
- Direction and interrogation messages should be consistent with the vocabulary commonly used by pilots
- Flexible syntax should be incorporated if possible
- System performance should remain constant for all phases of flight
- Failure mode provisions must be provided for and not degrade aircraft control
- Training on the recognition vocabulary should be able to be conducted in or out of the cockpit
- Training sessions should be self-prompting and inform the pilot if any words or phrases should be retrained
- The pilot should be able to update or retrain any word or phrase in or out of the cockpit

General Design Guidelines for Cockpit Voice Recognition Systems (sec. 2.4.2)

- Acceptable minimum accuracy and maximum error performance levels should be determined
- The operational ambient noise environment should be defined or estimated for normal, abnormal and emergency flight conditions
- Direction and interrogation messages associated with each task must be defined. From this a working vocabulary can be determined
- Message syntax should be noted and variable syntax combinations considered
- The vocabulary/messages should be divided into minimum subsets necessary for specific tasks or group of tasks so that the maximum active vocabulary sets can be defined
- Necessary response time should be estimated to compare with alternative methods
- If a recognition system controller is used, then first and second choice words and associated ratings must be available from the recognition system
- Pilot recognition vocabulary training provisions must be defined. This should include the ability to update individual words or phrases
- Interface hardware requirements must be defined and met
- Interface software requirements must be defined and met
- Environmental conditions (besides noise) must be defined
- Storage requirements for voice records must be considered and planned

Table 2.4.5 Summary of Design Guidelines for Cockpit Voice Synthesis Systems

Pilot-Based Design Guidelines for Cockpit Voice Synthesis Systems (sec. 2.3.6)

- Synthesis performance should be evaluated by pilots in the target cockpit environment with a representative message set
- Interferences with cockpit activities should be avoided

General Design Guidelines for Cockpit Voice Synthesis Systems (sec. 2.4.2)

- The operational ambient voice and noise environment should be defined or estimated for normal, abnormal and emergency flight conditions
- Acceptable minimum performance levels for intelligibility, voice quality and distinctiveness should be determined
- Voice messages associated with each task must be defined. From this a working vocabulary can be determined
- Presentation of voice messages to pilots must be planned
- The method of pilots to request information must be defined
- Interface hardware requirements must be defined
- Interface software requirements must be defined
- Environmental conditions (besides noise) must be defined
- Storage requirements for voice records must be considered and planned.

2.4.3 Candidate Cockpit and Simulator Voice Systems

Five generic voice recognition and synthesis systems are proposed below. The first three would be flight certified, per FAA standards, and used in existing to next-generation commercial aircraft. Two more systems are proposed that use commercially available equipment and would be useful in simulators (cockpit and control) and test aircraft. All the systems were designed with the guidelines that were specified in Sections 2.3.6 and 2.4.2 in mind. The cockpit and simulator systems suggested to interact with these voice systems are the ones that received the best benefit ratings in Section 2.4.1.

Voice System 1

This system is technically feasible today and is based on the AFTI/F-16 systems. Although AFTI-type systems are limited in usefulness, they are the only flight-qualified system available today. With isolated word recognition and limited recognition and synthesis vocabularies, such a system would not be able to interact with more than one or two cockpit systems. Also, the unidirectional data bus (ARINC-429) that is used on late-model commercial aircraft is not convenient for adding extra systems. It may be that the best way to incorporate this system for experimentation purposes would be to have it replace an existing control unit. A likely candidate would be the control display unit (CDU) where it could program and interrogate the navigation, inertial reference, and performance management systems. These voice tasks are rated as having high benefit to cockpit operations and, therefore, would be good first application of an interactive voice system. This voice system would meet all equipment requirements per FAA standards.

- Voice System 1: Recognition Specifications, Minimum Performance
  - > The system will operate with isolated word recognition but without word-spotting capability.
  - > The total vocabulary shall be ⩾100 words/phrases.
  - > Each word/phrase can have a maximum length of about 2 sec.
  - > It will be possible to program at least 20 vocabulary subdivisions (nodes).
  - > Each node shall be able to accommodate up to 25 words/phrases that will all be active/available at one time (active vocabulary).
  - > The overall recognition accuracy shall be ⩾95% in an ambient cockpit noise environment of ⩽90 dB.
  - > The system shall not experience a combined substitution and false acceptance error rate of greater than 0.5%.
  - > The pilot should be able to train on the vocabulary in the cockpit or at a ground training facility.
  - > During training the pilot will have to say each vocabulary word/phrase four to five times.
  - > The pilot shall have the option to update/retrain any single word without having to update the entire vocabulary.
  - > The voice data should be stored on some type of portable data storage module (DSM) that the pilots can take from the training facility to the aircraft and download their files to the voice system.
  - > The recognition system should operate equally well with boom or oxygen mask microphones.
  - > The pilot shall get feedback from the system via a preentry display.
  - > A yoke-mounted switch should be used to key the recognition system. When the switch is not depressed, no messages should be accepted.

- Voice System 1: Synthesis Specifications
  - > The digitally compressed method of voice synthesis/playback shall be used as being most desirable.
  - > There shall be about 200 sec of speech available to use.
  - > The speech synthesis data record shall be stored on the data storage module also.
  - > The voice model shall be selected using the guidelines from Sections 2.3.6 and 2.4.1.

- Voice System 1: Systems Controller Specifications
  > The controller shall have the necessary software to mimic a CDU to control the navigation, inertial reference, and performance management systems.
  > The controller shall accommodate only fixed syntax messages.
  > All software for general operation, training, and failure modes shall comply with the Section 2.3.6 and 2.4.2 guidelines.

- Voice System 1: General Specifications
  > At least one set of ARINC-429 interfaces (one input and one output per set) shall be available for connecting with the cockpit systems.
  > The system as a whole shall meet ARINC electrical and environmental specifications for commercial aircraft systems.
  > The system packaging shall meet ATR specifications.

- Voice System 1: Estimated Small Quantities Cost - $50K to 100K
  > This price estimate includes the recognition, synthesis, controller, and interface equipment. Also included is the necessary development software. Actual operations software is not included.

Voice System 2

This system will be technically feasible to be certified as a flight-quality system in one to three years. It is comparable to some of the better commercial-grade recognition and synthesis systems available today. This system will incorporate connected word recognition and phoneme-type synthesis. Therefore, it will approach the voice system requirements indicated in Sections 2.4.1 and 2.4.2. Because commercial aircraft will be using the ARINC-429 data bus to interface systems in this time period, it will be assumed that this system will require this type of interface also. A high-speed bidirectional data bus would be more useful and could be substituted for the ARINC-429 interfaces. In addition to the CDU (that system 1 interfaced with), this system can also interact with one or more of the following future concept systems: an advanced alerting system, integrated communications radio controller, integrated navigation radio controller, and integrated systems management system.

- Voice System 2: Recognition Specifications
  > The system shall operate with connected word recognition and have word-spotting capability.
  > The total vocabulary shall be ⩾300 words/phrases.
  > Each word/phrase can have a maximum length of about 2 sec.
  > It will be possible to program at least 40 vocabulary subdivisions (nodes).
  > Each node shall be able to accommodate up to 50 words/phrases, and these shall be active/available as a subset at any one time (active vocabulary).
  > The overall recognition accuracy shall be ⩾98% in an ambient cockpit noise environment of ⩽90 dB.
  > The system shall not experience a combined substitution and false acceptance error rate of greater than 0.1%.
  > The pilot shall be able to train on the vocabulary in the cockpit or at a ground training facility.
  > During training the pilot will have to say each vocabulary word/phrase no more than three to four times.
  > The pilot shall have the option to update/retrain any single word without having to update the entire vocabulary.

- > The voice data should be stored on some type of portable data storage module (DSM) that the pilots can take from the training facility to the aircraft and download their files to the voice system.
- > The recognition system should operate equally well with boom or oxygen mask microphones.
- > The pilot should get feedback from the system via a preentry display.
- > A yoke-mounted switch should be used to key the recognition system. When the switch is not depressed no messages should be accepted.

- Voice System 2: Synthesis Specifications
  - > Phoneme-type voice synthesis will be used.
  - > The total vocabulary will be at least 10,000 words/phrases.
  - > The voice model will be selected using the guidelines from Sections 2.3.6 and 2.4.1.

- Voice System 2: Systems Controller Specifications
  - > The controller will be able to control and interrogate the systems noted above.
  - > Limited syntax flexibility will be permitted.
  - > All software for general operation, training, and failure modes will comply with the Section 2.3.6 and 2.4.2 guidelines.

- Voice System 2: General Specifications
  - > Several sets of ARINC-429 interfaces (one input and one output per set) shall be available for connecting with the cockpit systems.
  - > The system as a whole shall meet ARINC electrical and environmental specifications for commercial aircraft systems.
  - > The system packaging shall meet ATR specifications.

- Voice System 2: Estimated Small Quantities Cost -$80K to 150K in 1 to 3 Years
  - > This price estimate includes the recognition, synthesis, controller, and interface equipment. Also included is the necessary development software. Actual operations software is not included.

Voice System 3

A system as described below should be technically feasible to a certifiable system in three to five years. It is assumed that by this time voice systems could be an integrated part of an advanced cockpit that would have many systems tied together via a high-speed bidirectional data bus. This voice system would probably have direct ties with a multifunction control display unit (MFCDU) so that the two systems would parallel each other's activities. The MFCDU display could serve as a preentry display and scratch pad for the voice system as well as presenting system information. All other systems could be interfaced through the MFCDU.

- Voice System 3: Recognition Specifications
  - > The system shall operate with connected word recognition and have word-spotting capability.
  - > The total vocabulary shall be ⩾500 words/phrases.
  - > Each word/phrase can have a maximum length of about 2 sec.
  - > It will be possible to program at least 70 vocabulary subdivisions (nodes).
  - > Each node shall be able to accommodate up to 60 words/phrases, and these will all be active/available at one time (active vocabulary).
  - > The overall recognition accuracy shall be ⩾99% in an ambient cockpit noise environment of ⩽90 dB.

- The system shall not experience a combined substitution and false acceptance error rate of greater than 0.05%.
- The pilot shall be able to train on the vocabulary in the cockpit or at a ground training facility.
- During training the pilot shall have to say each vocabulary word/phrase three to four times.
- The pilot shall have the option to update/retrain any single word without having to update the entire vocabulary.
- This system shall have limited speaker adaptability, i.e., the system will be able to update vocabulary words automatically if it has difficulty recognizing any of them.
- The voice data should be stored on some type of portable data storage module (DSM) that the pilots can take from the training facility to the aircraft and download their files to the voice system.
- The recognition system should operate equally well with boom or oxygen mask microphones.
- The pilot should get feedback from the system via a preentry and/or the MFCDU display.
- A yoke-mounted switch should be used to key the recognition system. When not depressed no messages should be accepted.

- Voice System 3: Synthesis Specifications
  - Phoneme-type voice synthesis will be used.
  - The total vocabulary will be at least 20,000 words.
  - The voice model will be selected using the guidelines from Sections 2.3.6 and 2.4.1.

- Voice System 3: Systems Controller Specifications
  - The controller will be able to control and interrogate the systems noted above and utilize an expert system data base.
  - Flexible syntax will be available.
  - All software for general operation, training, and failure modes will comply with the Section 2.3.6 and 2.4.2 guidelines.

- Voice System 3: General Specifications
  - The interface with the MFCDU and any other cockpit systems will be via a high-speed bidirectional data bus.
  - The system as a whole shall meet ARINC electrical and environmental specifications for commercial aircraft systems.
  - The system packaging shall meet ATR specifications.

- Voice System 3: Estimated Small Quantities Cost - $100K to 200K in 3 to 5 Years
  - This price estimate includes the recognition, synthesis, controller, and interface equipment. Also included is the necessary development software. Actual operations software is not included.

Voice System 4

This system is intended for use in cockpit simulators or test aircraft and is equivalent to current commercial-grade voice recognition and synthesis systems. Several interface options are available such as 16-bit bidirectional parallel, RS-232 or 422 serial or even ARINC-429 data buses. Depending on the cockpit device of interest, this system could use one or more of these interfaces. If this equipment is to be used in a motion-based simulator or test aircraft, some modifications to the systems may be necessary, e.g., securing printed circuit boards and providing uninterruptible power supplies. Also, if used in a test aircraft, normal safety precautions for experimental testing should be observed.

- Voice System 4: Recognition Specifications
  - > Per system 2.

- Voice System 4: Synthesis Specifications
  - > Per system 2.

- Voice System 4: Systems Controller Specifications
  - > Per system 2.

- Voice System 4: General Specifications
  - > This system may use 16-bit parallel, serial, or ARINC-429 interfaces.
  - > No ARINC or ATR regulations need be met except that system use on an aircraft must not interfere with normal flight operations and must not be capable of disabling any aircraft systems.

- Voice System 4: Estimated Small Quantities Cost - $15K to 25K
  - > This price estimate includes the recognition, synthesis, controller, and interface equipment. Also included is the necessary development software. Actual operations software is not included.
  - > As noted above, voice systems 4 and 5 could be built from existing commercial equipment. For example, the voice recognition controller and recognition system could be a Texas Instruments professional computer (with a speech command option and development software) or an IBM personal computer (with a Votan VPL-2000 board and development software). The phoneme synthesizer function could be handled by a DECtalk™ system (Digital Equipment Corp.) or a Call Text 5000 system (Speech Plus, Inc.). These systems are described in Appendix A.

Voice System 5

This system is similar to system 4 in that it assumes currently available commercial-grade equipment will be used. It will be used for controlling a simulator and not interacting with cockpit equipment. Section 2.2.3 discusses some of these applications. Only one bidirectional 16-bit data bus will be required.

2.5 Task 5: Comparison of Results With NASA Study of 1995 Transport

The final portion of this study was to compare the task 4 conclusions on cockpit voice applications to those proposed in a recently completed NASA sponsored study (ref. 31), titled "Crew Systems and Flight Station Concepts for a 1995 Transport Aircraft," by George A. Sexton, April 1983.

The "1995" study discussed voice systems possibilities, but did not define voice system guidelines that should be followed, nor did it recommend what methods of voice recognition or synthesis should be used. Also, there were no priority ratings for the voice tasks. Two lists of possible voice applications were provided, one for baseline cockpit voice requirements and the other for additional tasks intended for postbaseline voice recognition and synthesis requirements and to relate them to the proposed voice applications noted in task 4. An overview of the 1995 aircraft concept and the method of incorporating a voice system into baseline and postbaseline designs is also provided below.

2.5.1 Overview of the Proposed 1995 Transport Cockpit

A model for a 1995 transport aircraft was proposed in this study; although an entire aircraft was defined, the primary emphasis was on crew systems and the flight station. The flight deck features digital electronic fly-by-wire/light flight and thrust control systems, head-up displays, touch panel control for aircraft functional systems, voice command and response systems, and 1990 onboard air traffic control systems. Its two-pilot flight deck has numerous multifunction displays and control panels conveniently arranged on a glare shield, main instrument panel, and full-cockpit-width desktop. Some controls are located on an overhead panel, a center console, and two side pedestals, but most tasks would be monitored and controlled from the glare shield, main instrument panel, and desktop work area.

The flight deck design includes five monitoring and/or control modes that would have optional voice control. These modes are (1) the flight management computer (FMC) control/display units (CDU), (2) the combined communications/navigation radio frequency entry keyboard and display, (3) the front panel multifunction display system which includes five 13-in color CRT displays, three of which have touch panel overlays, (4) a guidance and control panel, and (5) the head-up displays. The voice command and control system was included in the cockpit system design so that the pilots would have a hands-off option for monitoring and controlling the aircraft subsystems. Table 2.5.1 lists the task requirements for the voice command and response system in the 1995 baseline design. A number of additional tasks were listed as possible requirements for a postbaseline design and these are included on Table 2.5.2.

All the flight deck subsystems are linked by redundant high-speed, bidirectional data buses and monitored/controlled by the FMCs. It is via these buses that the voice command and response system (and in turn the pilots) can interrogate, control, and respond to the other cockpit subsystems.

2.5.2 1995 Flight Management Computer (FMC) Control/Display Unit (CDU)

Two redundant FMCs are at the heart of the 1995 cockpit. The FMCs, via the bidirectional data buses, would be linked to all the aircraft flight systems, sensors, displays, and data input/output devices. The pilots monitor and control the various aircraft systems with the FMCs. The pilots would interact directly with the FMCs with two CDUs. (One CDU would be located in front of each pilot.)

The proposed CDUs have four major components: a 12-row by 40-character plasma display, a fixed key data entry typewriter (QWERTY-type) keyboard, a touch-panel faceplate for the display, and a group of low-profile membrane switches around the lower half of the display. With the CDU a pilot could call up 63 different formatted pages of information about aircraft location and systems status, plus permit control of navigation, performance management, communications, and electrical power controller systems.

Table 2.5.1 Baseline Voice Requirements for 1995 Cockpit

Voice Recognition (Control) Requirements
• Programming, interrogation, and data entry tasks (1) Call up of control/display unit (CDU) pages (2) Entering navigational waypoints (3) Call up of formats/information on the three center multifunction displays (4) Communications/navigation radio tuning and frequency entry • Selecting modes and switching tasks (1) Rain removal control, e.g., on/off (2) Landing lights control, e.g., on/off
Voice Synthesis (Response) Requirements
• Automatic voice messages (1) Barometric altitude alerts (2) Radar altitude alerts (3) Airspeed readouts (4) Time-critical messages • Positive action collision avoidance commands • Windshear and windshear/go-around alerts • Ground proximity warning system messages • Landing gear warnings • Pilot selectable voice messages (1) Readout of mode-S messages (2) Readout of ARINC Communications Addressing and Reporting System (ACARS) messages (3) Readout of advisory, caution and warning system messages (4) Takeoff and landing data information readout (5) Echo of voice entries

Table 2.5.2 Postbaseline Voice Requirements for 1995 Cockpit

Systems Which May Require Voice Interface (Recognition)
• Programming, interrogation, and data entry tasks (1) CDU—scratch pad entry of text messages (2) Global positioning system—waypoint entry (3) Guidance and control panel—altitude, flight-path angle, track, course, Mach, and indicated airspeed settings (4) Checklist display—item checkoff (5) Entry of mode-S transponder messages (6) Entry of ACARS messages • Selecting modes and switching tasks (1) Radar panel—mode select (2) Navigation display panel—range, display symbology selections (3) Comm/nav—transponder ident, active/standby transfer (4) Head-up display—declutter modes, on/off (5) Cabin advisory—seat belts and no smoking, on/off (6) Landing gear and brakes—autobrake, on/off (7) Lights—taxi lights, on/off (8) Systems display—all switches
Systems That May Require Voice Output (Synthesis)
• Pilot selectable voice messages (1) Echo of above items (2) Checklist—readout (3) Systems display—quantity readouts, status readouts (4) Engine display—parameter readouts

The proposed baseline voice command (recognition) system would allow the pilots to call up any of the formatted pages directly without having to sort through several to get the desired one. Voice commands could also be used to enter navigation way points instead of typing them in. The 12th CDU display line was designated as a scratch pad for the voice system when it works with the CDU.

A postbaseline voice application would be for entering text messages into the CDU scratch pad. The messages could be used for flight planning, mode-S transponder transmissions, or ACARS transmissions.

Comment: The programming and interrogating CDUs with voice were rated in task 4 as having a high benefit, but connected or continuous word recognition was recommended as necessary for this rating.

2.5.3 1995 Integrated Communications/Navigation Systems

The 1995 cockpit would have all the aircraft radios tuned and monitored with the integrated communications/navigation (ICN) system. Primary control and display of radio frequencies would be from a centrally located frequency entry keyboard and display panel. Redundant/optional controls would be available through one of the CDU formatted displays and the voice command system.

The radio frequency entry keyboard has its own preentry display for use with the keyboard or the voice entry modes. Once a frequency has been entered and verified, it is transferred to the desired active or standby radio display on the frequency display.

Comment: The use of voice for entering radio frequencies was rated in task 4 as having some benefit. An integrated comm/nav system such as this would simplify the adapting of a voice controlled system over a nonintegrated system as is presently found on commercial aircraft.

2.5.4 1995 Front Panel Multifunction Display System

The multifunction display system (MDS) features five 13-in diagonal color cathode ray tubes (CRT) and associated symbol generators. The CRTs are located on the main instrument panel with the two outside displays centered on the pilots' centerline positions. In normal operation the two outside CRTs would display primary and secondary flight information. The three center CRTs could be configured to display a number of information formats. The formats available, although not all simultaneously, include engine power, engine status, cockpit display of weather information, instrument approach information (Jeppeson charts), advisory, caution and warning information, cockpit display of traffic information, obstacle clearance detector information, operational emergency checklists, functional aircraft systems schematics, and menus for selecting desired display formats and information.

Should a CRT fail, priority information could be shifted to any other CRT. As noted above, the three center CRTs would have touch-panel overlays so pilots could interact directly with displayed information, e.g., checklists, schematics.

Baseline voice command requirements for the 1995 cockpit would give the pilots the option to select formats and information on the three center CRTs. This voice option would be a parallel task to the touch panel menu selections. A proposed postbaseline voice task would allow the pilots to check off items on the checklists.

Comment: The use of voice recognition to page, request information from, and interact with displayed information on multipurpose/multifunction displays was noted in task 4 as having a potentially high benefit.

2.5.5 1995: Additional Voice Recognition Requirements

All but two of the 1995 baseline and postbaseline voice recognition (control) requirements for programming, interrogation, and data entry have been noted above. One of those not noted yet is the use of voice to enter way points into a global positioning system (postbaseline task). Although a global (satellite) positioning system was identified as a possible future system in the task 2 subsystems overview, it was not included as a line item for any of the ratings. The use of voice recognition for tuning conventional navigation radios was found to be beneficial in the task 4 ratings. Using voice to tune a global positioning system should be comparable and therefore beneficial also.

The other voice recognition task proposed would be to enter data into the guidance and control panel (autopilot). Entering data into the autopilot with voice was rated as beneficial in task 4, although not as highly as programming and interrogation tasks. Designating this task for postbaseline introduction, after more beneficial tasks have been implemented, would appear to be a good choice.

Two voice-controlled mode select and switching tasks were proposed for the baseline cockpit design. Eight of these tasks were proposed for the postbaseline upgrades.

Comment: The task 4 benefits ratings indicate that switching and mode selecting tasks would be of uncertain benefit to the pilots. The uncertainty is due to the belief that using voice for switching and mode selecting tasks might be of little utility or advantage to the pilots. Some of these tasks would be useful in a hands/eyes overload situation, e.g., declutter control for the HUD, but in most cases the advantage would be questionable.

2.5.6 Voice Synthesis Requirements

Automatic and pilot selectable voice synthesis (response) requirements were proposed for the 1995 baseline cockpit. Four types of automatic voice synthesis messages were defined: barometric altitude alerts, radar altitude alerts, airspeed readouts, and time critical messages.

Comment: The automatic presentation of voice messages, other than time critical messages, is contrary to the task 4 guidelines, which were based on an FAA report on aircraft alerting systems design guidelines, Berson. et al, 1981. It is therefore recommended that the use of automatic presentation of voice messages, other than time critical, be carefully examined. Several pilot selectable voice tasks were proposed for the baseline and postbaseline designs. Many of the pilot selectable are also listed as potential applications in task 4 and are rated highly beneficial if pilot selectable.

2.5.7 Summary of Comparison of Present Results and 1995 Concepts

In most instances, the proposed uses of voice recognition and synthesis in the 1995 aircraft agree with the proposed high benefit applications recommended in task 4.

The proposed method for integrating voice into the 1995 cockpit is consistent with the task 4 guidelines on three important points. First, the 1995 design has voice control as an option to the pilot and always working in parallel with at least one other data entry device. Second, preentry displays are used. Lastly, most of the voice response messages were pilot selectable.

There are two points of difference between the two studies. One is the proposed use of voice control for switching in the baseline configuration. Using voice control for switching may be useful in some cases, but there are tasks with more potential, e.g., programming and interrogation tasks, that would seem to have a higher benefit to the pilots. The other inconsistency is the proposed automatic presentation of three nontime critical alerts. Guidelines for caution-warning standardization, as well as consideration of flight deck noise pollution, suggest that these possibilities be reconsidered.

3.0 Summary and Conclusions

This report presents a number of potentially beneficial applications of voice systems in commercial aircraft cockpits and simulators and proposes the necessary human factor design guidelines for their implementation. Six specific objectives were identified for the study to examine and report upon: (1) survey state-of-the-art voice technology and forecast developments in the next five years, (2) define and appraise the practicality of candidate applications in commercial aircraft cockpits and simulators, (3) identify the applications' suitability for actual operations, (4) develop a hierarchy of the applications based on their benefits and tradeoffs, (5) generate general specifications for aircraft and simulation voice systems, and (6) compare the results with those proposed in another recently completed NASA study of future flight decks.

The stated objectives were examined and answered by subdividing the study into five tasks, which are summarized below.

Task 1 reviewed state-of-the-art voice recognition and synthesis technology to establish a baseline. The review involved surveying literature, voice systems manufacturers, expert and general users, and an academic research center. From the surveys, information was gathered about what capabilities, performances, and limitations can be realistically expected for voice systems today and in the near future. Also, insights were gained into aircraft voice applications now in study and planned.

Three types of recognition systems are commercially available: isolated (all), connected (a few), and continuous (three) word recognition. The highest performance for all the recognition systems reviewed was in the isolated mode of recognition. Performance measures include the percentage of vocabulary words accurately recognized, percentage of words incorrectly substituted, number of vocabulary words recognized at one time, and the ability of a system to correctly recognize words in high ambient noise. Limited vocabulary (about 100 words) isolated word recognition systems have been demonstrated in and are available for high noise and g_n-force environments of fighter aircraft. Target military cockpits identified in surveys included F-16s, F-18s, KC-135s, and Blackhawk helicopters. Similar systems could be configured for commercial aircraft cockpits. Commercially available connected mode recognition systems should be available for commercial cockpits in one to three years.

The two most common types of voice synthesis systems in use today use digitally-compressed voice playback and phoneme-based voice synthesis. Good quality digital compressed systems are currently available that meet military and commercial flight requirements. Good quality phoneme-type voice synthesis systems are commercially available now and could be adapted to commercial flight requirements in one to three years.

Task 2 first reviewed commercial aircraft pilots' management requirements to identify possible voice recognition and synthesis applications, then rated their probable utility. Both existing and expected near-term aircraft subsystems were considered. A number of potential applications were identified for each subsystem, and requirements/constraints for each were specified, e.g., vocabulary size, necessary accuracy, and task frequency.

The potential cockpit recognition applications can be subdivided into five general types of tasks: programming, interrogating, data entry, switching/mode selection, and continuous control. The programming and interrogating of complex cockpit systems, e.g., control display unit, with voice will be advantageous to pilots if a connected word recognition system with at least 98% recognition accuracy is used. Data entry tasks (e.g., tuning radios) were generally found to be advantageous to the pilots even though they normally could be accomplished with isolated word recognition systems with at least 98% recognition accuracy. Switch/mode selection-type tasks were workable with isolated mode recognizers, but generally they gave no advantage to the pilots in performing the task. Finally, using voice to control continuous tasks, e.g., primary attitude controls, was estimated to be disadvantageous even with a connected word recognition system.

The potential cockpit synthesis applications that were identified included alerting messages, system status reports, secondary flight instrument data callout (airspeed and altitude), and playback of digitized communications messages. Basic alerts and data callout functions can be performed adequately with digitally compressed voice playback, but when large vocabularies are required the phoneme-type systems are recommended.

Task 3 examined the environmental considerations a designer must use when planning to integrate a voice system into a cockpit. Of the environmental requirements considered, ambient noise was found to be the most critical constraint. Without noise canceling techniques, most recognition systems experience high error rates when the ambient noise level exceeds 85-95 dB. Excellent noise canceling techniques are available, but they require signal preprocessors or use some of the capabilities of the speech recognition processor.

After considering the list of potential voice applications, some specific voice recognition and synthesis performance considerations were proposed. The performance of voice recognition systems was found to depend on recognition accuracy, substitution and insertion error rates, active vocabulary size, and message syntaxing capabilities. Intelligibility, distinctiveness from competing voice and noise sources, and voice model quality must be considered for synthesis performance.

Finally, a list of pilot-based design guidelines was proposed. The guidelines encompassed the following criteria: establishing the recognition performance in the target cockpit; determining what influence the pilot's knowledge and proficiency will have on system performance; providing corrective actions and clear feedback; allowing flexible message syntaxing; failure mode provisions; and providing adequate guidance through task operations so that operating manuals are not needed.

Task 4 used the information generated in the first three tasks to propose a benefits hierarchical rating of all the potential voice tasks. The proposed voice recognition applications with the most benefit to commercial cockpits involve programming and interrogating complex cockpit systems. Another high benefit area is the control of aircraft simulators. Data entry applications are the next most likely use of voice recognition, followed by switching and mode selection applications.

Several voice synthesis applications received high benefit ratings in task 4. One important reason for this is that all proposed voice messages would be at the pilot's request, except for time critical alerts. Pilot selectable voice messages prevent unnecessary cockpit chatter and missed messages.

As an extension to the task 3 pilot-based guidelines, a set of general-purpose design guidelines were identified. The general-purpose design guidelines took into account voice system performance criteria and hardware requirements.

Using both the guidelines and expected cockpit applications, five generic voice systems (recognition and synthesis) were proposed. The first system has limited capability (100 word vocabulary) and performance (≥95%) isolated word recognition system with digitally condensed voice response. It is based on systems developed for the Air Force AFTI/F-16 program and could be available for use on commercial aircraft, but it would not be available (certified) for one to three years. It is based on existing technology available in commercially available systems, and it would offer connected word recognition and phoneme-based voice synthesis. Both recognition and synthesis vocabularies would be larger than the first system. A third system would be available for commercial aircraft in three to five years. It would offer improved recognition accuracy, larger vocabularies, connected word recognition, and an integrated expert (artificial intelligence) computer system. The fourth system would control cockpit systems, but in simulators or test aircraft. It would use a commercially available connected mode recognizer and phoneme-type synthesizer. The last system proposed would be identical to the fourth in capability, but it would control the simulator systems (weather, emergencies, etc.) and not interact with the cockpit systems.

Task 5 concludes this report by comparing the recommendations for applying voice systems to commercial cockpits with those suggested in another NASA-sponsored study. This second NASA study developed a 1995 commercial jet air craft concept for an all-electronic cockpit. The all-electronic cockpit offers voice as an option to the pilots for several cockpit operations. The voice recognition applications in the baseline 1995 cockpit agreed with present study results in most cases by identifying similar programming, interrogating, and data entry tasks. Voice synthesis applications identified were also similar. The 1995 cockpit synthesis uses were divided into pilot selectable (for most) and automatic (for special alerts).

In conclusion, the importance of voice systems, at least initially, is less to replace existing cockpit hardware than to provide pilots an option in situations of high hands/eyes overload. More extensive research should be conducted to determine if this option would help pilots manage the cockpit systems more efficiently and safely.

Voice recognition systems would be best utilized to program and interrogate some of the more complex systems on the flight deck, e.g., select menus and specific aircraft data on a control display unit (CDU), and then enter new information if desired. Another programming application would be to prepare and send a message by data link to the ground (ATC or company office) or another aircraft. The second most likely application of recognition would be for entering data such as tuning radios or setting navigational way points. Voice can also be used to select switch or mode positions, but this was not estimated to be the best use of voice recognition.

Technology for voice synthesis (playback) systems is several years ahead of recognition systems, and they are already in limited use in several commercial aircraft models today.

The implementation of voice systems in the flight deck is just as important as the voice capabilities themselves. Before a system is selected, careful examination of the flight deck requirements and guidelines is necessary. If a voice system cannot benefit the operational requirement, it is best not to use it.

References

1. American Voice Input/Output Society, "Proceeding of the Voice Data Entry Systems Applications Conference 1983," American Voice Input/Output Society and Abbott Laboratories; Chicago, September 1983.

2. Armstrong, J. W. and Poock, G. K., "The Effects of Concurrent Motor Tasking on Performance of a Voice Recognition Systems," Thesis, Naval Postgraduate School; September 1980.

3. Armstrong, J. W. and Poock, G. K., "Effects of Operator Mental Loading on Voice Recognition System Performance," Naval Postgraduate School, NPS55-81-016; August 1981.

4. Armstrong, J. W. and Poock, G. K., "Effect of Task Duration on Voice Recognition System Performance," Naval Postgraduate School, NPS55-81-017; September 1981.

5. Batchellor, M. P., and Poock, G. K., "Investigation of Parameters Affecting Voice Recognition Systems in C3 Systems," Thesis, Naval Postgraduate School; March 1981.

6. Berson, B. L.; Po-Chedley, D. A.; Boucek, G. P.; Hanson, D. C.; Leffler, M. F.; and Wasson, R. L.; "Aircraft Alerting Systems Standardization Study, Vol. II - Aircraft Alerting System Design Guidelines," final report to FAA, DOT/FAA/RD-81/38,11, Vol. II; January, 1981.

7. Boucek, G.; Po Chedley, D. A.; Berson, B. L.; Hanson, D. C.; Leffler, M. F.; White, R. W.; "Aircraft Alerting Systems Standardization Study, Vol. I - Candidate System Validation and Time-Critical Display Evaluation," final report to FAA, DOT/FAA/RD-81/38,1; January 1981.

8. Clark, J. E., "A Review of Techniques for Speech Synthesis," Journal of Electrical and Electronics Engineering, Australia, IE Aust. & IREE Aust., Vol. 1, No. 1; March 1981.

9. Connolly, D. W., "Voice Data Entry in Air Traffic Control," Federal Aviation Administration, National Aviation Facilities Experimental Center, Report No. FAA-NA-79-20, August 1979.

10. Doddington, G. R. and Schalk, T. B., "Speech Recognition: Turning Theory to Practice," IEEE Spectrum, 1981, P. 26-32.

11. Davenport, D. D., unpublished Boeing study, 1980.

12. Feuge, R. L. and Geer, C. W., "Integrated Applications of Automated Speech Technology - Final Report," Office of Naval Research, ONR-CR213-158-1AF; February 1978.

13. French, B. A., and Poock, G. K., "Some Effects of Stress on Users of a Voice Recognition System - A Preliminary Inquiry," Thesis, Naval Postgraduate School; March 1983.

14. "Guidelines for Performance Assessment of Speech Recognizers," National Bureau of Standards, Dr. David Pallet, publication pending, May 1984.

15. Hecker, H. L., et al, "Manifestations of Task Induced Stress in the Acoustic Speech Signal," Journal of the Acoustical Society of America, v. 44, P. 993-1001; October 1968.

16. Jay, G. T., and Poock, G. K., "An Experiment in Voice Data Entry for Imagery Interpretation Reporting," Thesis, Naval Postgraduate School; March 1981.

17. Lea, Wayne A., "Commercial Speech Recognizers: A Comparison of Their Specifications and Reported Performance Capabilities," technical report to Naval Training Equipment Center, NAVTRAEQIPCEN 79-C-0141-4; September 1980.

18. Lea, W. A. and Woodard, J., "New Procedures for Comprehensive Assessment of Voice Entry Systems," AVIOS conference; September 1983.

19. McSorley, W. J., III, and Poock, G. K., "Using Voice Recognition Equipment to Run the Warfare Environmental Simulator (WEJ)," Thesis, Naval Postgraduate School; March 1981.

20. Mountford, S. J., Schwartz, J., and Graffunder, K., "Voice Interactive Prototype Automatic Target Screener," Honeywell Final Report, NASA, 83SRC28, Contract NAS2-11348; April 1983.

21. Parks, D. L., "Advanced Display Control Concepts for Power Plant Operation," IEEE 1981 Symposium on Nuclear Power Systems, San Francisco, October 1981.

22. Neil, D. E., and Andreason, J., "Examination of Voice Recognition System to Function in a Bilingual Mode," Naval Postgraduate School, NPS55-81-004; March 1981.

23. Poock, G. K., "Experiments with Voice Input for Command and Control: Using Voice Input to Operate a Distributed Computer Network," Naval Postgraduate School, NPS55-80-016; April 1980.

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

25. Poock, G. K., Schwalm, N. D., and Roland, E. F., "Use of Voice Recognition Equipment with Stenographer Masks," Naval Postgraduate School, NPS55-82-028; October 1982.

26. Poock, G. K., and Roland, E. F., "Voice Recognition Accuracy: What Is Acceptable," Naval Postgraduate School, NPS55-82-030; November 1982.

27. Poock, G. K., Sihwalm, N. D., Masters, B. J., and Roland, E. F., "Trying for Speaker Independence in the Use of Speaker Dependent Voice Recognition Equipment," Naval Postgraduate School, NPS 55-82-032; December 1982.

28. Poock, G. K., Roland, E. F., and Schwalm, N. D., "Wearing Army Gas Masks While Talking to a Voice Recognition System," Naval Postgraduate School, NPS55-83-005; March 1983.

29. Poock, G. K., and Martin, B. J., "Voice Recognition with Naive versus Practiced Speakers," Naval Postgraduate School, NPS55-83-016; June 1983.

30. Scott, P. B., "Word Recognition," Rome Air Development Center Report RADC-TR-78-209; September 1978.

31. Sexton, G. A., "Crew Systems and Flight Stations Concepts for a 1995 Transport Aircraft," NASA Langley Research Center, CR-166068, April 1983.

32. Shochet, E., Quick, P., and Delemarre, L., "Flight Plan Filing by Speech Recognition," FAA Technical Center, Atlantic City, DOT/FAA/RD-82/39 and DOT/FAA/CT-81/64, AD A119679; July 1982.

33. Stokes, J. M., "VRAS System Guide," final report to NADC, Technical Report 1400.19-B, N62269-79-C-0235; February 1982.

34. Taggert, J. L., Wolfe, C. D., Neil, D. and Poock, G., "Voice Recognition as an Input Modality for the Tacco Preflight Data Insertion Task in the P-3C Aircraft," Thesis, Naval Postgraduate School; March 1981.

35. Yellen, H. W. and Poock, G. K., "A Preliminary Analysis of Human Factors Affecting the Recognition Accuracy of a Discrete Word Recognizer for C3 Systems," Thesis, Naval Postgraduate School; March 1983.

Appendix A. Voice Equipment Review

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
American Microsystems, Inc. 3800 Homestead Road Santa Clara, CA 95051 (408) 246-0330	S360 LPC-10 speech synthesizer —	—	• Digitized and condensed • LPC-10 method	• 2K-bits speech at 2K-bits/sec max rate • Up to 32 words max	• Commercial product • 24-pin chip, CMOS with 20K bits on chip ROM
	S3620 LPC-10 speech synthesizer —	—	• Digitized and condensed • LPC-10 method		• Commercial product • 22 pin CMOS chip
E-Systems, Inc. P.O. Box 226118 Dallas, TX 75266 (214) 272-0515	CV3670/A —	—	• LPC-10 and channel vocoders	• 2.4K-bits/sec digitizing	• For JTIDS data link of voice • Commercial and military products
General Digital Corp. 700 Bornside Avenue East Hartford, CN 06108, (203) 528-9041	GDX— Speech-TI —	—	• Digitized • With TI LPC-10 tech	• Basic—206 industrial words • ROM space for 206 additional • Can take LPC code from host	• Multibus expansion module—per Intel SBX expansion bus specification • Commercial product
General Instrument 600 West John Street Hicksville, NY 11802 (516) 733-3107	SP1000 —	• User-programmed 8-stage LPC • Independent and dependent capabilities	• Digitized and condensed • 10-stage LPC	• User design dependent	• 28-pin chip—user to incorporate in own system • Commercial product
	SPO256A-AL2 —	—	• Digitized and condensed	• Allophones on ROM, no limit	• Chip • Commercial product

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Covox Company 675-D Conger Street Eugene, OR 97402 (503) 342-1271	Voice Master $120	—	• Digitized and condensed	• Up to 150 words	• Module to plug to Commodore 64 PC • Commercial product
Digital Equipment Corp. 146 Main Street Maynard, MA 01754	DECTalk™ $4K	—	Phoneme + text-to-speech	• Basic vocabulary of >20K words • 120-135 words per sec	• Standalone with RS232C I/O • Commercial product
Digital Sound Corp. 2030 Alameda Padre Serra Santa Barbara, CA 93103 (805) 963-8951	DSC-200™ $20K	• High quality digitizing system for rec. or syn. work research	—	• 32K bytes/sec typical • 1.6M bytes/sec maximum	• Interfaces with PDP or VAX system • Requires host control • Commercial product
Dragon Systems, Inc. Chapel Bridge Park 55 Chapel Street Newton, MA 02158 (617) 527-0372	Dragon Mark II (Priced for OEM)	• Discrete • Dependent or independent	—	• Memory dependent	• Software for small computer on host (i.e., Apple II) • Commercial product • Software + OEM design support
	Dragon Mark (OEM System)	• Continuous • Dependent or independent	— Limited	• Memory dependent	• Design uses small computer (IBM PC) on host + custom board • Prototype • Software + OEM support

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
ICS Electronics Corp 1620 Zanker Road San Jose, CA 95112 (408) 298-4844	4800 —	—	• Digitized and condensed	• Basic 300 words • Space for custom vocabulary	• Standalone • Controlled via IEEE-488 bus • Commercial product
INFOVOX AB Box 121, 5-18212 Danderyd, Sweden (468) 753-3460	RA 101 $3.5K	• Discrete • Speaker dependent • Syntax available	—	• Basic 48 • Option to 3066 but not real-time	• Standalone • I/O via RS232C • Commercial product
	SA 101 $3.5K	—	• Phoneme • Text-to-speech	• Rate and pitch control	• Standalone • I/O via RS232C • Commercial product
	RA 101/PC • For OEM $500/each in 500 unit quantity	• Discrete • Speaker dependent available	— • Syntax	• Basic 48 with upload/download	• Board level for IBM PC • OEM to provide host (IBM PC) software • Commercial product due June 1984
	SA 101/PC • For OEM $750/each in 500 unit quantity	—	• Phoneme • Text-to-speech	• Rate and pitch control	• Board level for IBM PC • OEM to provide software • Commercial product due June 1984

Appendix A. **Voice Equipment Review** (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Intel Corp. 3065 Bowers Avenue Santa Clara, CA 95051 (408) 987-8080	iSBC 570 —	• Discrete • Syntaxing available • Speaker dependent	• Digitized and condensed	• Memory dependent vocabulary base	• Board (iSBC 576) + control unit • Requires Intel development system • Commercial product • Extensive software development package
	iSBC 576 —	• Discrete • Syntaxing available • Speaker dependent	• Digitized and condensed	• Basic 200 word rec.	• Board level • I/O via Intel iSBC multibus or RS232 to other host systems • Commercial product
	iSBC 577 —	• Discrete • Syntaxing available • Speaker dependent	• Digitized and condensed	• Depends on user design	• Chip set (8) • Prototype
Interstate Electronics Corp. 1001 East Ball Road Anaheim, CA 92803 (714)	VRT 101 Family —	• Discrete • Syntaxing available • Speaker dependent	—	• 100 word resident • Up/download with disk or host	• Standalone • 2 - RS232C I/O • Commercial product • CP/M based operating system
	VRT 300 —	• Discrete • Syntax available • Speaker dependent	—	• 200 word resident • Up/download with host	• Board level for DEC VT100 and others • Commercial product • Resident firmware
	VRT 200 —	• Discrete • Syntax available • Speaker dependent	—	• 100 word resident • Up/download with host	• Board level for ADM 3A and 5 terminals • Commercial product • Resident firmware

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Interstate Electronics Corp. (Cont.)	SYS 300	• Discrete • Syntax available • Speaker dependent	—	• 100 word resident • Up/download with host	• Standalone • 2—RS232C I/O • Commercial product • Resident firmware
	VRM 102 —	• Discrete • Syntax available • Speaker dependent	—	• 100 word resident • Up/download with host	• Multibus board • Commercial product • Resident firmware
	VTM 150 —	—	• Phoneme based	• 500 program and 1K user define	• Multibus board • Serial and parallel • Commercial product • Resident firmware
	VRC 100-2 —	• Discrete • Speaker dependent • Syntax available	—	• User design, up to 200 addressable	• 2-chip set • Commercial product • Firmware on ROM
ITT Def. Electronics Corp 10060 Carroll Canyon Rd. San Diego, CA 92131 (619) 578-3080	Voice command system (VCS)	• Discrete • Speaker dependent • Syntax—up to 25 nodes of 25 words	• Digitized and condensed	• 100 words rec. • 125 seconds speech syn.	• Standalone • I/O via RS232 or military standard 1553 • Military qualified • Resident firmware
	VCS for IBM PC —	• Discrete and connected • Syntaxing	—	TBD	• Board level for IBM PC • Commercial prototype

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Key Tronic P.O. Box 14687 Spokane, WA 99214 (509) 928-8000	5152 speech recognition keyboard $1.5K	• Discrete • Speaker dependent • Syntaxing—9 subdivisions	—	• 100 words resident • Up/download with host	• Keybaord that plugs into IBM PC • Commercial product • Training and operating software included
Lear Siegler, Inc. Instrument Division 4141 Eastern Ave. SE Grand Rapids, MI 49507 (616) 241-7000	Voice-controlled interactive device —	• Discrete • Speaker dependent • Syntaxing available	—	• 100 words nominally • Data storage on removable module	• Standalone unit • I/O via Military Standard 1553B • Military qualified • Training and operating software included
Micromint Inc. 561 Willow Avenue Cedarhurst, NY 11516 (516) 374-6793	Microvox —	—	Phoneme • Text to speech	• 64 phonemes	• Standalone • I/O via parallel or RS-232C • 1000-3000-character buffer • Commercial product • Resident firmware
Microvoice Systems Corp. 33362 Peralta Drive Suite 5 Laguna Hills, CA 92653	Voiceboard $375 to $820	—	• Digitized and condensed	• Up to 8 minutes • 4096 word basic • User-defined words available	• Single board for OEMs • Commercial product • Resident firmware
National Semiconductor Corp 2900 Semiconductor Drive Santa Clara, CA 95051 (408) 737-5000	MM54104 Digitalker —	—	• Digitized and condensed with LPC method	• Several hundred words available • For user's circuit • Will make custom vocabularies	• Chip level • Commercial product

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
NEC America, Inc. 532 Broad Hollow Rd Melville, NY 11747 (516) 752-9700	DP-200 CSR $9K	• Discrete and connected • Speaker dependent • Syntaxing available	—	• Basic 50 words connected speech • Options=150 words connected and up to 500 words discrete mode	• Standalone system • I/O with host via RS232C, RS422, GPIB and 8-bit parallel • Commercial product • Training and application software included
	SR-100 $2K	• Discrete • Speaker dependent • Syntaxing by host computer	—	• Basic 150 words • Up/download with host	• Host controlled • I/O with host via RS232C • Commercial product • User developed software
OKI Semiconductor Corp 1333 Lawrence Expressway Suite 401 Santa Clara, CA 95051 (408) 848-4840	SAS-1 Real-Voice Memory Processor $9K	—	• Digitized and condensed	• 128-sec maximum voice input at 4 kHz and 64-sec at 8 kHz • 64-sec buffer for editing	• Standalone for recording own messages and transfer to phone • Commercial product
Sanyo Semiconductor Corp. 7 Pearl Court Allendale, NJ 07401 (201) 825-8080	LC 8100 —	—	• Digitized and condensed • 10-stage filter	• User defined • 28 seconds on chip • Up to 28 minutes on external ROM	• CMOS chip, 28 pin • Commercial product

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
SCI Systems Inc. 8600 So. Memorial Pkwy. P.O. Box 4000 Huntsville, AL 35802 (205) 882-4800	Voice Control Unit I —	• Discrete • Speaker dependent • Syntaxing available	• Digitized and condensed • LPC method	• 100 word rec. • 100 word syn. • Up/download with host	• Standalone system • I/O with host via military standard 1553B • Military qualified • Resident training and operations software
	Voice Control Unit II —	• Discrete and connected + word spotting • Speaker dependent • Syntaxing available	• Digitized and condensed • LPC method	• 150 word rec. at one time • Up/download with host	• Standalone system • I/O with host via military standard 1553B • Military qualified • Resident training and operations software
Scott Instruments 1111 Willow Springs Drive Denton, TX 76201 (817) 387-9514	VET-2 $800	• Discrete, speaker dependent	—	• 40 words • 1.5-sec duration per word • Up/download with host	• For use with Apple II and Franklin Ace 1000 computers • Board level and comes with microphone • Commercial product
Speech Plus, Incorp 461 North Bernardo Mountain View, CA 94043 (415) 964-7023	Call text 5000 —	—	• Text to speech • Phoneme based	• Rate=50 to 250 words/minute	• Board for IBM PC • I/O—RS232C and telephony • Commercial product • Applications software per MS-DOS

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Speech Plus, Incorp. (cont.)	SP1020A $2.5K	—	• Digitized and condensed with LPC	• 2200 bits/sec standard • EPROM for up to 36 seconds • Words recorded by vendor	• Standalone • I/O with host via RS232C • Board versions available • Commercial product • Resident firmware
	PR2000 $3.5K	—	• Text to speech • Phoneme based	• Rate=50 to 250 words/minute • Size dependent on host memory • Words recorded by vendor	• Multibus board • I/O via RS232C • Commercial product • Resident firmware
Speech Systems, Inc. 18356 Oxnard Street Tarzana, CA 91356 (818) 881-0885	— —	• Continuous • Phoneme based • Limited training, system adapts to user	—	• Now 200 word vocabulary • Goal of 5000 word	• Standalone system • Aim '85 commercial product • Prototype • Software designed for dictation
Street Electronics Corp 1140 Mark Avenue Carpinteria, CA 93013 (805) 684-4593	Echo Speech Board $200	—	• Text to speech • Phoneme based	• Basic with speech • Option of 90 seconds custom speech	• Board level for phoneme codes Apple II • Commercial product • Resident firmware
	Echo GP —	—	• Text to speech • Phoneme based	• Basic with phoneme codes • Option of 90 seconds custom speech	• Standalone • I/O via RS232C

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Street Electronics Corp. (Cont.)	Echo PC —	—	• Text to speech • Phoneme based	• Basic with phoneme codes • Option of 90 seconds custom speech	• Standalone for use with IBM PC
Texas Instrument, Inc. Speech Applications P.O. Box 226015, M/S 394 Dallas, TX 75266 (214) 995-6571	Speech Command System $2.6K	• Connected word spotting discrete • Speaker dependent • Syntaxing for 9 nodes	• Digitized and condensed per TI-LPC	• Syn word size dependent on memory • Syn/rec at 2400 bits/sec • 50 words per vocabulary • Up/download	• Board for TI PC • Commercial product • Training and applications software on disk
	Voice interactive system —	• Connected word spotting + discrete • Speaker dependent • Syntaxing for 9 nodes	• Digitized and condensed per TI-LPC	• Syn word size dependent on memory • Syn/rec at 2400 bits/sec • 50 words per vocabulary • Up/download • Module for voice records	• Based on TI speech command system • Standalone system • Military qualified • Resident firmware
	TMS 5100 voice system vocabulary processor —	—	• Digitized and condensed per TI-LPC	• Number of words depends on design • Up to 30 minute speech can be addressed	• Chip • Large available custom work • Commercial product

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Threshold Technology Inc. 1829 Underwood Blvd. Delran, NJ 08075 (609) 461-4200	CSR 1000 —	• Continuous, word-in-context • Speaker dependent	—	• 1000 word total • ? active at one time	• Host controlled • I/O with host via RS232C • Board level available—Intel multibus format • Prototype • User design software
Verbex Two Oak Park Bedford, MA 01730 (617) 275-5160	Model 3000 $17.9K	• Connected and discrete • Speaker dependent	—	• 120 standard size and option to 360 words	• Standalone unit • I/O with host via RS232C • Need SPADS system to develop programs • Commercial product • Training and operating software
	Model 3000 SPADS $32K	• Connected and discrete • Speaker dependent	—	• 120 standard size and option to 360 words	• Standalone unit • Development system with terminal and hard disk storage • Commercial product • Training and operating software • Applications development software

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Votan 4487 Technology Drive Fremont, CA 94538 (415) 490-7600	VX Series V1000-V6000 —	• Discrete • Syntaxing available • Standard speaker dependent with limited speaker independent option	• Digitized and condensed record and playback	• Up to 255 word recognition vocabulary, total • 60-70 words active at one time • Up/download data with host	• Standalone and board products available • I/O to host via RS232C • Local or remote control • Commercial product • Training operating and development software available
	V8000 series —	• Discrete • Syntaxing available • Standard speaker dependent	• Digitized and condensed record and playback	• Up to 255 word recognition vocabulary, total • 60-70 words active • Up/download data with host	• Standalone designed for IBM PC host • Includes IBM PC • Commercial product • Training operating and development software available • Software interactive with IBM PC
	VPC-2000 $2.5K	• Continuous + word spotting + discrete • Speaker dependent • Syntaxing available	• Digitized and condensed record and playback	• 75 words active • Up/download data with host • Synthesis rate 4K to 14.4K bits/sec	• Board for IBM PC • Includes telephone management • Prototype—due May 1984 • Training and operating and development software

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Votan (cont.)	VSP-1000 $2.5K	• Continuous word spotting and discrete • Speaker independent option	• Digitized and condensed record and playback	• Up to 250 words discrete rec • Up to 75 words discrete rec • Up to 75 words • Continuous rec • Synthesis rate 600-1.8K bits/sec • Up/download data with host	• Multibus board • Prototype • Resident firmware • User design software
	VTR-6000 $4K	• Continuous and word spotting and discrete • Speaker dependent • Syntaxing available	• Digitized and condensed record and playback	• Up to 250 words discrete rec • Up to 75 words continuous rate • Synthesis rate 600 to 1.8K bits/sec • Up/download data with host	• Standalone • I/O with host via RS232C • Prototype • User design software • Resident firmware
Votrax 500 Stephenson Highway Troy, MI 48084 (313) 588-2050	Type-n-talk $300	—	• Phoneme text to speech	• Data rate = 70-100 bits/sec	• Standalone • Receive data from host via RS232C • Commercial product • Resident firmware
	SC-01 $40.00	—	• Phoneme text to speech	• Data rate = 70 to 100 bits/sec	• Chip • User designed circuit • Commercial product • Resident firmware

Appendix A. Voice Equipment Review (Continued)

Manufacturer and Location	Model and Price	Recognition Capabilities	Synthesis Capabilities	Vocabulary Size Basic and Options	Availability, Packaging, and Software Support
Votrax (Cont.)	VS-6 $3.6K	—	• Phoneme coding principles	• 61 phonemes + 4 inflection levels available • 8-bit command word	• Standalone • Receives data from host via RS232C or 8-bit parallel • Commercial product • Resident firmware
	ML-I —	—	• Phoneme coding principles	• 122 phonemes and 8 inflection levels (pitch) 4 phoneme rates (duration) • 12-bit command word	• Standalone • Receives data from host via RS232C or parallel • Multilingual • Commercial product • Resident firmware

Appendix B
Representative Bibliography, Post-1980

Advisory Group for Aerospace Research and Development, "Advanced Avionics and the Military Aircraft Man/Machine Interface," AGARD-CP-329 ISBN-92-835-0315-4, Neuilly-sur-Seine, France, July 1982.

Aimthikul, Y., "Computer-Generated Speech," DE81-004672 ORNL/TM-896, W-7405-ENG-26, NTS SAP: HC A05/MF A01, Oak Ridge National Laboratory, Tennessee, December 1981.

Anders, R., Kryway, J. T., "Advanced Speech Technology in Fighter Cockpits - A New Perspective on Issues and Applications," A83-11083 01-01, Logicon, Inc., San Diego, California, NAECON 1982: Proceedings of Natl. Aerospace & Electronics Conf., Vol. 1, p. 351-355, Dayton, Ohio, May 18-20, 1982.

Anderson, J. C., "An Extremely Low-Cost Computer Voice Response System," C81014740, BYTE, Vol. 6, No. 2, 36-43, Mimic Electronics, Acton, Massachusetts, February 1981.

Anderson, J. C., "A Low-Cost Digital Voice Recognition and Response System," B81047155, C81032817, Telecommunications, Vol. 15, No. 6, 47-50, June 1981.

Armstrong, J. W., Poock, G. K., "Effect of Task Duration on Voice Recognition System Performance," AD-A107477 NPS55-81-016, N00039-80WR-09041, NTIS SAP: HC A04/MF A01, Naval Postgraduate School, Monterey, California, August 1981.

Baker, D., "Voice Synthesis Appliances: A Pragmatic Analysis," B83037536, IEEE Trans. Ind. Appl., Vol. IA-19, No. 1, 136-41, Percom Data Co., Inc., Dallas, Texas, Jan-Feb 1983.

Batchelor, Christopher L., "Noise Cancellation to Improve the Quality of LPC Processed Speech Degraded by Noise," AD-A115 539/9, Report No. AFIT/GE/EE/81D-10, Air Force Inst. of Tech., School of Engineering, Wright-Patterson AFB, Ohio, December 1981.

Bell, D. W., Becker, R. W., "Designing Experiments to Evaluate Speech I/O Devices and Applications," B83044564, Speech Technol., Vol. 1, No. 4, 70-5, 78-9, Bioengineering Res. Center, SRI International, Menlo Park, California, Jan-Feb 1983.

Bell, R., Bennett, M. E., Brown, W. E., "Direct Voice Input for the Cockpit Environment," B83004715, Flight Automation Res. Lab., Marconi Avionics Ltd., Rochester, England, AGARD Conf. Proceedings, No. 329, Advanced Avionics and the Man/Machine Interface (AGARD-CP-329) 13/1-7, Blackpool, England, April 26-29, 1982.

Blankenship, J., "Give Your Apple a Voice," C82036242, BYTE, Vol. 7, No. 5, p. 556-56, Devry Inst. of Technnol., Atlanta, Georgia, May 1982.

Bridle, J. S., "Connected-Word Recognition for Use in Military Systems," AD-B07207OL, AC/243-D/854 AC/243 (Panel III)D/218, NATO Defense Research Group, Brussels, Belgium, March 18, 1983.

Post-1980 (Continued)

Bui, Ngoc C., Monbaron, Jean J., Michel, Jean G., "An Integrated Voice Recognition System," E18307058897, IEEE J Solid State Circuits, Vol. SC-18 N. 1, p. 75-81, Asulab S. A., Neuchatel, Switzerland, February 1983.

Bui, Ngoc, C., Monbaron, J. J., Michel, J. G., IEEE Trans. Acoust. Speech & Signal Process., Vol. ASSP-31, No. 1, Pt. 2, p. 323-9, Asulab S. A., Neuchatel, Switzerland, February 1983.

Butler, F., Manaker, E., Obert-Thorn, W., Sherman, H., "Investigation of a Voice Synthesis System for the F-14 Aircraft," AD-B058705L ACT-81-001, N62269-79-C-0493, Grumman Aerospace Corp., Bethpage, New York, June 5, 1981.

Chi Foon Chan, Hoff, M. E., Jr., Nevard, P., Meemong Lee, "Architecture and Application of a Commercially Available Speech Recognition Board," B83044598, C83031036, Intel Corp., Santa Clara, California, Proceedings of ICASSP 83, IEEE International Converence on Acoustics, Speech and Signal Processing, Vol. 2, p. 495-8 Boston, Massachusetts, April 14-16, 1983.

Chi Foon Chan, Fink, D. F., "A Total Approach to Speech Recognition," Speech Technol., Vol. 1, No. 4, 116-17, Intel Corp., Santa Clara, California, Jan-Feb 1983.

Civil Engineering Lab. (Navy), "Voice Technology Efforts May Replace Keyboarding," PB82-970933, Port Hueneme, California, October 1982.

Cochran, D. J., Riley M. W., "Data Input by Voice," Comput. & Ind. Eng. (GB), Vol. 7, No. 2, 115-20, Dept of Industrial and Management Systems Eng., Univ. of Nebraska, Lincoln, Nebraska, 1983.

Condom, P., "Automatic Voice Recognition," B83027408, C83016908, Vol. 38, No. 1, p. 82-3, Interavia (Engl. Ed.), Switzerland, January 1983.

Cooke, N., "The Flight Testing and System Integration of Direct Voice Input Devices in the Civil Flight Deck," B83039926, Royal Aircraft Establ., Bedford, England, Colloquium on Digital Processing of Speech, 13/1-6, London, England, April 6, 1983.

Cooke, M. P., "A Speech Controlled Information-Retrieval System," NPL-DITC-15/83 ISSN-026205369, National Physical Lab., Teddington, England, January 1983.

Coster, J. R., Melocco, J. M., "Using Voice Control Onboard Combat Aircraft," RAE-TM-FS-412 BR80500, DCAF E010260, Royal Aircraft Establishment, Farnborough, England, June 15, 1981.

Costet, J. R., "Voice Techniques on Board Aircraft - An Experimental Approach," A83-11083-01-01, Crouzet, Division Aerospatial et Systemes, Valence, Drome, France NAECON 1982: Proceedings of Natl. Aerospace & Electronics Conf., Vol. 1, P 3390342, Dayton, Ohio, May 18-20, 1982.

Cotton, John C., McCauley, Michael E., "Voice Technology Design Guides for Navy Training Systems," AD-A129 371/1, Report No. HEL-TM-7-83, Human Engineering Lab., Aberdeen Proving Ground, Maryland, May 1983.

Cotton, J. C., McCauley, M. E., North, R. A., Strieb, M. I., "Development of Speech Input/Output Interfaces for Tactical Aircraft," Wright Aeronautical Laboratories, Wright-Patterson AFB, Ohio, Canyon Research Group, Inc., Westlake Village, California, July 1983.

Davies, E. B., Peckham, J. B., "New Developments in Cockpit-Human Interfaces," RAE-TM-FS-412 BR80500, DCAF E010260, Royal Aircraft Establishment, Farnborough, England, June 15, 1981.

Davis, Robert L., Sinnamon, James T., Cox, David L., "Voice Verification Upgrade," AD-A118 444/9, Report No. TI-08-82-07; RADC-TR-82-139, Texas Instruments, Inc., Dallas, Texas, June 1982.

Davis G., Stockton, G., "F-16 Voice Message System Study," A83-11083 01-01, General Dynamics Corp., Fort Worth, Texas, NAECON 1982: Proceedings of Natl. Aerospace & Electronics Conf., Vol. 1, p. 324-331, Dayton, Ohio, May 18-20, 1982.

DeJackmo, M. L., "Taking a Closer Look at Speech Recognition and Computer Voice Synthesis," Gov. Data Syst., Vol. 10, No. 1, p. 23-4, International Resource Dev., Norwalk, Connecticut, Jan-Feb 1981.

Demauduitduplissix, L., "Speech Recognition and Synthesis," VAU/82/101, DCAF E090905, NTIS SAP: HC A07/MF A01., Engins Matra, Velizy France, 1982.

Desipio, R. G., Fry, E., "Avionics System Plays 'Ask and Tell' With Its Operator," B83046545, C83029971, Speech Technol., Vol. 1, No. 4, 89-93, Techol. Directorate Naval Air Development Center, Warminster, Pennsylvania Jan-Feb 1983.

Dusek, L., Schalk, T. B., McMahan, M., "Voice Recognition Joins Speech on Programmable Board," B83038305, C82027826, Electronics, Vol. 56, No. 8, 128-32, Texas Instruments, Inc., Dallas, Texas, April 21, 1983.

Emerson, S. T., "Voice Response Systems-Technology to the Rescue for Business Users," Speech Technol., Vol. 1, No. 4, p. 99-101, 103, Periphonics Corp., Bohemia, New York, Jan-Feb 1983.

French, Brian Arthur, "Some Effects of Stress on Users of a Voice Recognition System: A Preliminary Inquiry," AD-A128 559/2, Naval Postgraduate School, Monterey, California, March 1983.

Frehse, R., Geppert, R., Gierloff, R., Kuhn, M. H., Ney, H., Piptrowski, H., Tomaschewski, H., "Automatic Recognition of Cooperative Speakers," BM-FT-DV80-009 ISSN-0170-9011, Philips Zentrallaboratorium G.m.b.H., Hamburg, West Germany, December 1980.

Gill, S. P., "VOTAN Voice Technology Adapted for Use in Military Environments," B83046360, C83031051, VOTAN, Fremont, California, Publ: IEEE, New York, IEEE Proceedings, National Aerospace and Electronics Conference, NAECON 1983, Vol. 1, p. 413-15, Dayton, Ohio, May 17-19, 1983.

Gooding, M. J., "Voice ATLAS," A83-10726 01-59, Proceedings of Autotestcon Conf., Honeywell, Inc., Minneapolis, Minnesota, October 19-21, 1981.

Gruber, J., Strawczynski, L., "Subjective Effects of Variable Delay and Speech Loss in Dynamically Managed Voice Systems," B83019674, Bell-Northern Res., Ottawa, Ontario, Canada, IEEE Global Telecommunications Conf., 1329-33, Vol. 3, Miami, Florida, November 29-December 2, 1982.

Post-1980 (Continued)

Gupta, V., Mermelstein, P., "Speaker-Independent Word Recognition Techniques for Control of a Voice Messaging System," B82010676, C82007674, Bell-Northern Res., Nuns' Island, Quebec, Canada, Proceedings of the 7th Canadian Man-Computer Communications Conference, 233-8, Waterloo, Ontario, Canada, June 10-12, 1981.

Handwerker, J., "AFSATCOM-Compatible 75 B/S Voice System," B83019697, C83015095, IEEE, Military Commun. Conf. Progress in Spread Spectrum Communications, 21/3/1-4, vol. 1, Mitre Corp, Bedford, Massachusetts, 1982.

Harkins, A. M., Joels, K. M., "An On-Site Delphi at Princeton '81," A82-35601 17-12, Proceedings of the Fifth Space Manufacturing Conference, Princeton, New Jersey, May 18-21, 1981.

Holmgren, John E., "Applying Automatic Speech Recognition to Telephone Services," EI8306048387, IEEE Commun. Mag. Vol. 20, N. 6, p. 31-34, American Bell, Inc., Systems Planning Dept., November 1982.

Holmaren, J. E., "Toward Bell System Applications of Automatic Speech Recognition," The Bell System Technical Journal, Vol. 62, No. 6, July-August 1983.

Houck, J. A., "Evaluation of a Voice Recognition System for the MOTAS Pseudo Pilot Station Function," NASA-TM-84487 NAS 1.15:84487, National Aeronautics and Space Administration, Langley Research Center, Hampton, Virginia, April 1982.

Hunter, J. Calvin, "Time Axis Analysis of Gravity Distorted Speech," AD-A115 540/7, Report No. AFIT/GE/EE/81D-27, Air Force Inst. of Tech., School of Engineering, Wright-Patterson AFB, Ohio, December 1981.

Institut fuer Angewandte Geodaesie, "Reports on Cartography and Geodesy," ISSN-0469-4236, NTIS SAP: HC A05/MF A01, Frankfurt am Main, West Germany, 1980.

Kayser, Jeffrey A., "The Correlation Between Subjective and Objective Measures of Coded Speech Quality and Intelligibility Following Noise Corruption," AD-A115 515/9, Report No. AFIT/GE/EE/81D-30, Air Force Inst. of Tech., School of Engineering, Wright-Patterson AFB, Ohio, December 1981.

Koekebakker, J., "Computer Voice Response/Recognition. Making Progress Towards a Difficult Objective," B82002296, C82004349, Vol. 13, No. 5, 265-7, 29, 31, Can. Datasyst., Canada, May 1981.

Lea, W. A., "Selecting the Best Speech Recognizer for the Job," B83044560, Speech Technol., Vol 1, No. 4, 10-22, 27-9, Speech Sci. Publications, Santa Barbara, California, Jan-Feb 1983.

Lightstone, P. C., Davidson, W. M., "Voice Synthesis Application," DE82015482, Report No. SAND-82-0415, Sandia National Labs., Albuquerque, New Mexico, April 1982.

Lundquist, E., "Voice-Input Systems Make Inroads Into Industrial Applications," C83010479, Mini-Micro Syst., Vol. 15, No. 10, 165-72, October 1982.

Post-1980 (Continued)

Malkin, Frank J., "Effects on Computer Recognition of Speech When Speaking for Navy Training Systems," AD-A129 145/9, Report No. CRG-TR-82-006; NAVTRAEQUIPC-80-C-0057-1, Canyon Research Group, Inc., Westlake Village, California, March 1983.

Manaker, Elliott, "Pilot Ability to Understand Synthetic Voice and Radio Voice When Received Simultaneously," AD-A119 137/8, Report No. ACT-R-82-01, Grumman Aerospace Corp., Bethpage, New York, June 1982.

Martin, B. J., Roland, E. F., "The Effect of Feedback to Users of Voice Recognition Equipment," AD-A127223, NPS55-83-003, MIPR-TB-024, Naval Postgraduate School, Monterey, California, February 1983.

Mayer, Mitchell S., "Electronic Voice Communications Improvements for Army Aircraft," AD-A119 427/3, Report No. USAAVRADCOM-TR-81-E-1, Army Avionics Research and Development Activity, Fort Monmouth, New Jersey, August 1982.

McComb, G., "Low Cost Voice Synthesis: All Ears for Type'N'Talk," C82011941, Creative Comput., Vol. 7, No. 9, p. 148-51, September 1981.

Meidinger, C. L., "Automatic Secure Voice Communications (AUTOSEVOCOM) Technical Evaluation Program Report for Fort Ritchie, Maryland, Command Operations Center Follow-on Operational Test and Evaluation Report," AD-B073020L, REPT-1815-AT-83-021, Test and Evaluation Squadron (1815th), Wright-Patterson AFB, Ohio, March 29, 1983.

Mod. Off. Procedures, "Voice Programming-A Lucky Accident Makes It Reality," C83005898, Vol. 27, No. 9, p. 154, September 1982.

Moore, T. J., Nixon, C. W., McKinley, R. L., "Comparative Intelligibility of Speech Materials Processed by Standard Air Force Voice Communication Systems in the Presence of Simulated Cockpit Noise," SAP HC A09/MF A01, AGARD, N81-31449 22-32, Aerospace Medical Research Labs., Wright-Patterson AFB, Ohio, June 1981.

Mountford, S. J., Schwartz, J., Graffunder, K., "Voice Interactive Prototype Automatic Target Screener," Final Report on Contract No. NAS2-11348 to Army Aeromechanics Lab, Moffet Field, California, Systems and Research Center, Minneapolis, Minnesota, April 1983.

Munro, Allen, Towne, Douglas M., Cody, James A., Abramowski, Harry, "Techniques for Computer-Based Training of Air Intercept Decision Making Skills," AD-A125 373/1, Report No. TR-101, University of Southern California, Behavioral Technology Labs, Redondo Beach, California, October 1982.

National Technical Information Service, "Speech Intelligibility, 1979-February 1983 (Citations from the NTIS Data Base," PB83-804286, Springfield, Virginia, March 1983.

National Technical Information Service, "Speech Recognition by Computer. Octobver 1981-1982 (A Bibliography with Abstracts)," PB83-802835, Springfield, Virginia, February 1983.

Nitta, T., Murata, T., Tsuboi, H., Takeda, K., Kawada, T., Watanabe, S., "Development of Japanese Voice-Activated Word Processor Using Isolated Monosyllable Recognition," C83008540, C83005917, Toshiba Res. and Dev. Center, Kawasaki, Japan, Proceedings of ICASSP 82, IEEE International Conf. on Acoustics, Speech and Signal Processing, Vol. 2, p. 871-4, Paris, France, May 3-5, 1982.

North, R., Lea W., "Application of Advanced Speech Technology in Manned Penetration Bombers," AD-A119 274/9, Report No. 81SRC 57; AFWAL-TR-82-3004, Honeywell Systems and Research Center, Minneapolis, Minnesota, March 1982.

O'Neil, E. F., "Voice Entry: Terminals You Can Talk To," C83005899, Data Commun., Vol. 11, No. 10, 133-4, 137-8, 141, Interstate Electronics Corp., Anaheim, California, October 1982.

Poock, Gary K., Martin, B. Jay, "Voice Recognition Performance with Naive versus Practiced Speakers," AD-A130 155/5, Report No. NPS55-83-016, Naval Postgraduate School, Monterey, California, June 1983.

Poock, G. K., Roland, E. F., "Voice Recognition Vocabulary Lists for the Army's TACFIRE System," AD-129 975/9, Report No. NPS55-83-001, Naval Postgraduate School, Monterey, California, January 1983.

Poock, G. K., Roland, E. F., Schwalm, N. D., "Wearing Army Gas Masks While Taling to a Voice Recognition System," AD-A129 951/9, Report No. NPS55-83-005, Naval Postgraduate School, Monterey, California, January 1983.

Poock, Gary K., Schwalm, Norman D., Martin, B. Jay, Roland, Ellen F., "Trying for Speaker Independence in the Use of Speaker Dependent Voice Recognition Equipment," AD-A129 938/7, Report No. NPS55-82-032, Naval Postgraduate School, Monterey, California, December 1982.

Poock, G. K., Roland, E. F., "Voice Recognition Accuracy: What is Acceptable," AD-A129 930/4, Report No. NPS55-82-030, Naval Postgraduate School, Monterey, California, November 1982.

Poock, G. K., Schwalm, N. D., Roland, E. F., "Use of Voice Recognition Equipment with Stenographer Masks," AD-A122 470/8, Report No. NPS55-82-028, Poock, G. K., "Voice Recognition Boosts Command Terminal Throughput," A83032553, B83019633, C83013796, Speech Technol., Vol. 1, No. 2, p. 36-9, U.S. Naval Postgraduate School, Monterey, California, April 1982.

Poock, G. K., Armstrong, J. W., "Effect of Operator Mental Loading on Voice Recognition System Performance," AD-A107477 NPS 55-81-016, N00039-80WR-09041, NTIS SAP": HC A04/MF A01, Naval Postgraduate School, Monterey, California, August 1981.

Poock, G. K., Martin, B. J., "Effects of Emotional and Perceptual Motor Stress on a Voice Recognition System's Accuracy: An Applied Investigation," Report No. NPS55-84-002, for Naval Electronics Systems Command, Washington, D.C., Naval Postgraduate School, Monterey, California, February 1984.

Pool, R. H., Kellett, H. G., "The Cockpit Voice Entry Trail-Where is it Going?," B83046359, C83031050, Collins Avionics Div., Rockwell International Corp., Cedar Rapids, Iowa, Publ: IEE, New York, IEEE Proceedings, National Aero space and Electronics Conference, NAECON 1983, 409-12 Vol. 1, No. 4, 89-93, Techol. Directorate Naval Air Development Center, Warminster, Pennsylvania, Jan-Feb 1983.

Rabiner, L. R., Levinson, S. E., Sondhi, M. M., "On the Use of Hidden Markov Models for Speaker-Independent Recognition of Isolated Words from a Medium-Size Vocabulary," AT&T Bell Laboratories Technical Journal, Vol. 63, No. 4, April 1984.

Reed, L. W., "Voice Interactive Systems Technology Avionics (VISTA) Program," B83004717, U.S. Army Avionics Res. and Dev. Activity, Fort Monmouth, New Jersey, AGARD Conf. Proceedings, No. 329, Advanced Avionics and the Military Aircraft Man/Machine Interface (AGARD-CP-329) 15/1-14, Blackpool, England, April 26-29, 1982.

Reinecke, M., "Voice Warning Systems: Some Experimental Evidence Concerning Applications," NTIS SAP: HC A09/MF A01, AGARD, N81-31449 p. 22-32, Air Force Inst. of Aviation Medicine, Manching, West Germany, June 1981.

Rosenberg, N., "The Votrax Type'N'Talk: Give Your Computer a Voice," C82042958, Interface Age, Vol. 7, No. 8, p. 98-100, August 1982.

Rubinchek, B., "Toward Standards for Speech I/O Systems," B83044562, Speech Techno., Vol. 1, No. 4, 40-2, Jan-Feb 1983.

Ruth, J. C., Godwin, A. M., Werkowitz, E. B., "Voice Interactive System Development Program (Military Aircraft Control)," B83004716, C83001311, General Dynamics Corp., Fort Worth, Texas, AGARD Conf. Proceedings, No. 329, Advanced Avionics and the Military Aircraft Man/Machine Interface (AGARD-CP-329) 14/1-10, Blackpool, England, April 26-29, 1982.

Schalk, T. B., Frantz, G. A., Woodson, L., "Voice Synthesis and Recognition," C83017955, Mini-Micro System, Vol. 15, No. 12, p. 145-60, Texas Instruments, Inc., Dallas, Texas, December 1982.

Schmidt-Nielsen, A., Everett, S. S., "A Conversational Test for Comparing Voice Systems Using Working Two-Way Communication Links," AD-A116054 NRL-8583, RR0210542, Naval Research Lab., Washington, D. C., June 11, 1982.

Stokes, James M., "VRAS System Guide," Technical Report 1400.19-B, submitted to Naval Air Development Center, Warminster, Pennsylvania, Code 6021, February 1982, Analytics, Willow Grove, Pennsylvania.

Street, M., Adler, B., "OEM Voice Board Breaks Price Barrier," B83044570, C83031027, Speech Technol., Vol. 1, No. 4, 112-15, Street Electronics Corp., Carpinteria, California, Jan-Feb 1983.

Teja, Edward R., "Board-Level Voice-Output Systems Help Products Speak for Themselves," E1811096457, EDN Vol. 26, No. 7, p. 45-59, April 1, 1981.

Teja, E. R., "Voice-Output Units Show Improvement in Speech Quality, Vocabulary, Price," B82060349, C82042951, EDN, Vol. 27, No. 11, p. 71-80, May 26, 1982.

Valdez, M. E. "Adding Voice to a Computer," C83013791, Micro - 6502/6809 J., No. 55, 69-73, December 1982.

Vonusa, R., Cupples, E., Steigerwald, S., Woodard, J., Nelson, J., "Application, Assessment and Enahncement of Speech Recognition for the Aircraft Environment," N83-18257 08-54, Rome Air Development Center, Griffiss AFB, New York, July 1982.

Wolf, J., Krasner, M., Karnofsky, K., Schwartz, R., Roucos, S., "Further Investigation of Probabilistic Methods for Text-Independent Speaker Identificatin," B83044609, Bolt Beranek and Newman Inc., Cambridge, Massachusetts, Proceedings of ICASSP 83, IEE International Conf. on Acoustics, Speech and Signal Processing, Vol. 2, p. 551-4 Boston, Massachusetts, April 14-16, 1983.

Withers, S. J., "Voice Control of an Interactive Simulation," C83022457, Simulation, Vol. 40, No. 1, p. 28-9, School of Industrial and Business Studies, University of Warwick, Coventry, England, January 1983.

Wood, R. P., Moore, C. A., "A Cost Effective Approach to Design Evaluation of Advanced System Display Switchology," NAECON 1982 Proceedings of Natl. Aerospace & Electronics Conf., Dayton, Ohio, May 18-20, 1982, Vol. 2, A83-11083 01-01, New York Inst. of Electrical and Electronics Engineers, Inc., p. 964-970, General Dynamics Corp., Fort Worth, Texas, 1982.

Yellen, Howard William, "Preliminary Analysis of Human Factors Affecting the REcognition Accuracy of a Discrete Word Recognizer for C3 Systems," AD-A128 546/9, Naval Postgraduate School, Monterey, California, March 1983.

Appendix C

Representative Bibliography, Through 1980

Aatre, V. K., Ataman, E., Wong, K. M., "Median Filtering – Nonlinear Signal Processing for Speech and Picture," Annual Alerton Conference on Communication, Control, and Computing, Proceedings, 17th, Monticello, Illinois, October 10-12, 1979, Nova Scotia Technical College, Halifax, Canada, 1979.

Anderson, R. R., Foschini, G. J., Gopinath, B., "A Queueing Model for a Hybrid Data Multiplexer," Bell System Technical Journal, Volume 58, February 1979, Bell Telephone Laboratories, Inc., Murray Hill, New Jersey, February 1979.

Apelwicz, T., "The Design of a Programmable Real Time ADM Voice Processor and the Formulation of the Autocorrelation Function for the LDM algorithm TLSP: Ph.D. Thesis," Microfilms Order No. 7902540, City Univ. of New York, New York, 1979.

Araseki, T., Itoh, Y., Ochiai, K., "A Microprocessor Echo Canceller – for Satellite Communication Networks," International Conference on Digital Satellite Communications, Proceedings, 4th, Montreal, Canada, October 23-25, 1978, Teleglobe Canada, 1979, p. 219-224, Nippon Electric Co., Ltd., Kawasakai, Japan, 1979.

Armstrong, J. W., "The Effects of Concurrent Motor Tasking on Performance of a Voice Recognition System," AD-A093 557/7, Naval Postgraduate School, Monterey, California, September 1980.

Ataman, E., Aatre, V. K., Wong, K. M., "A Fast Method for Real-Time Median Filtering – Algorithm for Signal and Image Processing," IEEE Transactions on Acoustics, Speech, and Signal Processing, Volume ASSP-28, August 1980, NCR, Waterloo, Ontario, Canada, Nova Scotia Technical College, Halifax, Canada, August 1980.

Ayala, I. L., "Aperiodic Noise Cancellation Using Adaptive Least Absolute Value Filters," Asilomar Conference on Circuits, Systems, and Computers (13th) Pacific Grove, California, November 5-7, 1979, Institute of Electrical and Electronics Engineers, Rockwell International Electronics Research Center, Anaheim, California, 1980.

Bahler, L., Markey, P., Moshier, S., "Intelligence Report Voice Input," Final Technical Report, AD-A062845 RADC-TR-78-239, Dialog Systems, Inc., Belmont, Massachusetts, November 1978.

Baird, R., "The SI-10000 Status and Identification System," Philips Telecommunications Mfg. Co. Ltd., Clayton, Australia, Philips Telecommunications, Rev. (Netherlands), Volume 36, No. 1 100-3, April 1978.

Barnwell, T. P. III, Voiers, W. D., "An Analysis of Objective Measures for User Acceptance of Voice Communications Systems," E21-659-78-TB-1, Georgia Inst. of Tech. Atlanta School of Electrical Engineering, Dynastat, Inc., Austin, Texas, September 1979.

Beek, B., Cupples, E. J., "Military Applications of Automatic Speech Recognition and Future Requirements," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Rome Air Development Center, Griffiss Air Force Base, New York, 1977.

Beek, B., Cupples, E. J., Heffron, M., "Speech Recognition as Related to Military Applications," Wescon Tech. Pap., Volume 22, West Electronics Show and Convention, Los Angeles, California, September 12-14, 1978, Publ: Electron Conv., El Segundo, California, 1978.

Berouti, M., Krasner, M., Makhoul, J., "Efficient Encoding and Decoding of Speech," Quarterly Progress Report, 1 Feb. - 30 Apr. 1980, AD-A085972, BBN-4384 QPR-2, Bolt, Beranek, and Newman, Inc., Cambridge, Massachusetts, May 1980.

Berouti, M., Makoul, J. Schwartz, R., Sorenson, J., "Speech Compression and Syntheses," Final Report 8 Jun. 1979 - 31 Jul. 1980, AD-A092578, BBN-4414 CNT#: F19628-78-C-0136 ARPA, Bolt, Beranek, and Newman, Inc., Cambridge, Massachusetts, October, 1980.

Bially, T., "Information Processing Techniques Program," Semiannual Technical Summary Report, 1 Apr. - 30 Sep. 1979, AD-A083195, ESD-TR-79-281, Lincoln Lab., Mass. Inst. of Tech., Lexington, September 30, 1979.

Bierfert, H., Kirstein, M., Lance, D., "Some Aspects of Evaluating the Performance of a Speech Recognition System in Real Applications," IEEE Int. Conf. Acoust. Speech Signal Process, ICASSP 80, Proc., Volume 1, p. 186-189, Denver, Colorado, April 9-11, 1980, Publ: IEEE (Cat No. 80CH1559-4), Piscataway, New Jersey.

Billi, R., Modena, G., "Artificial Intelligence and Man-Machine Dialogue by Voice," Atti del Congresso Annuale AICA 80 (Annual Conference AICA 80) 897-927, 1980, Part II 29-31, Technoprint, Bologna, Italy, October 1980.

Birch, J. N., "Comparison of Analog and Digital Transmissions for Multichannel Voice," International Telemetering Conference, Los Angeles, California, October 18-20, 1977, Instrument Society of America, 1977, p. 273-279, U. S. Department of Defense, Office of the Secretary of Defense, Washington D. C., 1977.

Blankenship, P. E., "A Review of Narrowband Speech Processing Techniques," Advanced Signal Processing Proceedings, Peebles, Scotland, September 18-21, 1979, Institution of Electrical Engineers, MIT, Lexington, Massachusetts, 1979.

Boll, S. F., Pulsipher, D. C., Done, W., Cox, B. V., Timothy, L. K., "Noise Suppression Methods for Robust Speech Processing TLSP: Semiannual Technical Report, 1 Apr. - 31 Sep. 1978," AD-A064694 UTEC-CSC-78-204, Utah Univ., Salt Lake City, December 1978.

Boll, Steven F., "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," IEEE Trans. Acoust. Speech Signal Process, Volume ASSP-27, No. 2, p. 113-120, Univ. of Utah, Salt Lake City, Utah, April 1979.

Boll, S. F., Pulsipher, D., Ravindra, H., "Noise Suppression Methods for Robust Speech Processing TLSP: Semiannual Technical Report, 1 Apr. - 30 Sep. 1979," AD-A079321 UTEC-CSC-79-163, Utah Univ., Salt Lake City, October 1979.

Bolt, R. A., "'Put-That-There' Voice and Gesture at the Graphics Interface," Comput. Graphics (ACM), Volume 14, No. 3, p. 262-70, SIGGRAPH '80 Conf. Proc. Annual conf. on Comput Graph and Interact Tech. 7th, Seattle, Washington, July 14-18, 1980, MIT, Cambridge, Massachusetts, July 1980.

Brodersen, R. W., Hurst, P. J., Allstot, D. J., Tsuda, T., "Switched- Capacitor Applications in Speech Processing," Advanced Signal Processing Proceedings, Peebles, Scotland, September 18-21, 1979, Institution of Electrical Engineers, California University, Berkeley, California, Fujitsu, Ltd., Kawasaki, Japan, 1979.

Bunge, E., "Automatic Forensic Speaker Recognition," Carnahan Conference on Crime Countermeasures Proceedings, Ky. Univ. Off. Res. Eng. Serv. Bull. UKY BU117, p. 41-45, Univ. of Kentucky, Lexington, Ky., May 16-18, 1979, Publ: Univ. of Ky. Coll. of Eng., ORES Publ.

Buslenko, N. P., Deev, V. V., Frolov, G. D., "A System of Automatic Voice communication with the Computer," Trans In: Program. and Comput. Software (USA), Volume 3, No. 3, 226-9, May-June 1977.

Buzo, A., Gray, A. H., Jr., Markel, J. D., Gray, R. M., "Speech Coding Based Upon Vector Quantization," IEEE Transactions on Acoustics, Speech, and Signal Processing, Volume ASSP-28, October 1980, Universidad Nacional Autonoma de Mexico, Mexico City, Mexico, Signal Technology, Inc., Santa Barbara, California, Stanford University, Stanford, California, October 1980.

Campbell, J. M., Saridis, G. N., "A Voice-Controlled Mechanical Arm for Immobilized Patients," ICASSP 79, 1979 IEEE International Conference on Acoustics, Speech and Signal Processing, 582-5, Washington, D. C., April 2-4, 1979, Publ: IEEE, New York.

Campanella, S. J., "Digital Speech Interpolation Techniques," NTC '78; National Telecommunications Conference, Birmingham, Alabama, December 3-6, 1978, Conference Record, Volume 1, Institute of Electrical and Electronics Engineers, Inc., 1978, p. 14.1.1-14.1.5, COMSAT Laboratories, Clarksburg, Maryland, 1978.

Cannings, R., "Speech Patterns and Aircrew Workload," Surv. of Methods to Assess Workload, AGARD, Royal Air Force Inst. of Aviation Medicine, Farnborough, England, August 1979.

Carrigan, B., "Applications of the Fast Fourier Transform," NTIS, Springfield, Va. September 1977-august, 1980.

Carson, L. M., "Lock Detector for Noise-Coded Signals," NASA Tech. Brief, Volume 4, No.3 Fall, 1979, p. 348, B79-10324, NOP-14435, PAA (Motorola, Inc.), April 1980.

Cheung, R. S., Winslow, R. L., "Subband Coder Study TLSP: Final Report," AD-A074427 AD-E100278, GTE Sylvania, Inc., Needham Heights, Massachusetts, July 2, 1979.

Cohen, Arnon, Graupe, Daniel, "Speech Recognition and Control System for the Severely Disabled," Journal of Biomedical Engineering, Volume 2, No.2, p. 97-107, Ben-Gurion Univ. of the Negev, Beer-Sheva, Israel, April 1980.

Cohen, J. R., Crystal, T. H. House, A. S., Neuburg, E. P., "Weighty Voices and Shaky Evidence, A Critique," J. Acoust. Soc. Am. (USA), Volume 68, No. 6, 1884-6, Department of Defense, Ft. Meade, Maryland, December 1980.

Cole, R. A., Jakimik, J., Cooper, W. E., "Perceptibility of Phonetic Features in Fluent Speech," J. Acoust. Soc. Am., Volume 64, No. 1, p. 44-56, Carnegie-Mellon Univ., Pittsburgh, Pennsylvania, July 1978.

Coler, C. R., Plummer, R. R., Huff, E. M., Hitchcock, M. H., "Automatic Speech Recognition Research at NASA-Ames Research Center," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., National Aeronautics and Space Administration, Ames Research Center, Moffett Field, California, 1977.

Congdon, K. P. Woodruff, K., "Voice Countermeasures Effectiveness Analysis," AFAL-TR-73-230, AD-912 892/8, NTIS, Systems Research Labs, Inc., Dayton, Ohio, July 1973.

Connolly, D. W., "Voice Data Entry in Air Traffic Control," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., National Aviation Facilities Experimental Center, Atlantic City, New Jersey, 1977.

Connolly, D. W., "Voice Data Entry in Air Traffic Control TLSP: Report, May 1975 - Dec. 1978," AD-A073670 FAA-NA-79-20, National Aviation Facilities Experimental Center, Atlantic City, New Jersey, August 1979.

Connolly, D. W., "Voice Data Entry in Air Traffic Control," FAA-NA-79-20, AD-A073 670/2ST, NTIS, National Aviation Facilities Experimental Center, Atlantic City, New Jersey, August 1979.

Curran, M. "Voice Integrated Systems," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Naval Air Development Center, Warminster, Pennsylvania, 1977.

Dante, H. M., Sarma, V. V. S., "Pattern Recognition Model of Voice-Based Personal Verification Systems for Forensic Applications," IEEE Trans. Syst. Man Cybern, Volume SMC-10, No. 9, p. 585-588, Indian Inst. of Sci., Bangalore, India, September 9, 1980.

David, G., "Digitalized Voice and Data Encoding Using the Same Bit Rate for Satellite Communications," Military Electronics Defence Expo '79; Proceedings of the Conference, Wiesbaden, West Germany, September 25-27, 1979, (A80-34151 13-04, Cointrin, Geneva, Switzerland, Interavia, S. A., 1979, p. 309-324, Telecommunications RadioElectriques et Telephoniques, Le Plessis-Robinson, Hauts-de-Seine, France, 1979).

Davidoff, A. E., "Voice-Actuated Controller for the Handicapped," Johns Hoplins APL Tech. Dig. (USA), Volume 1, No. 1, 57-8, January-March 1980.

Davies, E. B., Peckham, J. B., "New Developments in Cockpit Human Interfaces," International Council of the Aeronautical Sciences, Congress, Proceedings, 12th, Munich, West Germany, October 12-17, 1980, American Institute of Aeronautics and Astronautics, Inc., 1980, Royal Aircraft Establishment, Farnborough, Hants., England, 1980.

Davis, D. B., "Voice Data Entry Experiences Slow Market Acceptance," Minicomput. News (USA), Volume 6, No. 10, 16, 18 October 23, 1980.

Davis, R. L., Hydrick, B. M. Doddington, G. R., "Total Voice Speaker Verification," RADC-TR-78-260, Texas Instruments, Inc., Dallas, Texas, Rome Air Development Cetner, Griffiss AFB, New York, Jan. 1979.

Davis, R. L., "Application of Statistical Pattern Recognition to Speech Recognition and Voice Authentication," Paper IEEE comput Soc. Int. Conf., Compcon Spring '80, VLSI', New Architectural Horizons, San Francisco, California, February 25-28, 1980, IEEE, (Cat No. 80CH1491-0-C), p. 407-412, New York.

Dempsey, A. D., "A Not So Modest Proposal (Machine Reaching)," J. Educ. Technol. Syst. (USA), Volume 9, No. 1, 77-81, School of Education, Florida Internat. Univ., Miami, Florida, 1980-81.

Desjardins, R., "Natural Human-to-Machine Interface," Research and Technology Objectives and Plans Summary; Fiscal Year 1979 Research and Technology Program, NASA-TM-80035, p. 10, (W79-70522, 310-40-37), National Aeronautics and and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, 1979.

Doddington, G. R., "Speech Systems Research at Texas Instruments," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Texas Instruments, Inc., Dallas, Texas, 1977.

Doddington, G. R., "Voice Authentication Parameters and Usage," Wescon Tech. Paper, Volume 22, West Electron Show and Convention, Los Angeles, California, September 12-14, 1978, Publ: Electron Conv., El Segundo, California, 1978.

Dubes, R., Jain, A. K., Tosi, O. I., "Speaker Recognition Studies on Choral Speech," IEEE Proceedings of the 1979 Carnahan Conference on Crime Countermeasures, 165-71, Lexington, Kentucky, May 16-18, 1979, Publ: Univ. Kentucky, Lexington, Kentucky.

Dvorak, P., Ptacek, M., "Peripheral Unit for Speech Communication," Sdelovaci Tech. (Czechoslovakia), Volume 28, No. 5, 163-5, In Czech, May 1980.

Edwards, G. P., "A Speech/Speaker Recognition and Response System," ICASSP 80 Proceedings, IEEE International Conference on Acoustics, Speech and Signal Processing, 394-7, Part II, Denver, Colorado, April 9-11, 1980, IEEE, (Cat. No. 80CH1559-4), Piscataway, New Jersey.

Elchafie, C., "A System for the Automatic Recognition of the Speakers on a Minicomputer," Ph.D. Thesis, N80-18255, available from Univ. Microflims Order No. 8003040, DCAF F090913, NTIS, Paris-Sud Univ., Orsay, France, 1978.

Elchafie, C., Thomson CSF, Versailles, France, and Jackson, J. S., DeVore, R. W. (Editors), "A Real Time Automatic Speaker Recognition System on Mini-Computer," IEEE Proc. of the 1980 Carnahan Conference on Crime Countermeasures, Publ. by Univ. of Ky., Coll. of Eng., Off of Eng. Serv., Univ. Kentucky, Lexington, Ky., May 14-16, 1980.

Elster, R. S., "The Effects of Certain Background Noises on the Performance of a Voice Recognition System," AD-A106138, NPS 54-30-010, Naval Postgraduate School, Monterey, California, September 1980.

Englander, I., "Voice Response System Statistics Program, Operational Handbook," Final Research, Apr. 1978 - Dec. 1979, AD-A089109 TSC-FAA-80-9 FAA-RD-80-76, Transportation Systems Center, Cambridge, Massachusetts, June 1980.

Erber, N. P., De Filippo, C. L., "Voice/Mouth Synthesis and Tactual/Visual Perception of PA, BA, MA," J. Acoust. Soc. Am. (USA), Volume 64, No. 4, 1015-19, October 1978.

Fellbaum, K., "Methods of Digital Speech Signal Transmission," BLL-RDL-TRANS-3798-(7737,470), British Library Lending Div., Boston Spa, England, February 1980.

Ferber, L. A., "Voice Input/Output Capabilities at Perception Technology Corporation," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Perception Technology Corp., Winchester, Massachusetts, 1977.

Fette, B., Gibson, R., Greenwood, E., "Windowing Functions for the Average Magnitude Difference Function Pitch Extractor," ICASSP 80 Proceedings, IEEE International Conference on Acoustics, Speech and Signal Processing, 49-52, Part I, Denver Colorado, April 9-11, 1980, IEEE, New York.

Fetzner, W. N., "Definition of Problems of Persons in Sheltered Care Environments," Project Extension Report, Jan. - Dec. 1978, NASA-CR-160050 Wisconsin University, Madison, January 15, 1979.

Feuge, R. L., Geer, C. W., "Integrated Applications of Automated Speech Technology Final Report," ONR-CR213-158-1AF, to Office of Naval Research, Arlington, Virginia, Boeing Aerospace Company, Seattle, Washington, February 1978.

Finin, T., Goodman, B., Tennant, H., "Jets: Achieving Completeness Through Coverage and Closure," AD-A081957 WP-19, Illinois Univ. at Champaign-Urbana, February 8, 1979.

Fink, Daniel F., "Voice Input in Natural Language for Man-Machine Interfacing," E1810977865, Interstate Electron Corp., Anaheim, California, Publ: Control Eng., Barrington, Illinois, 1980, P. 103-107, Proc. Ann. Adv. Control Conf. 6th, Purdue University, West Lafayette, Indiana, April 28-30, 1980.

Flanagan, J. L., "Digital Voice Techniques for Transmission, Storage and Automatic Recognition," Proceedings of Compcon Fall '78, Computer Communications Networks, Washington, D. C., September 5-8, 1978, Publ: IEEE, New York.

Flanagan, J. L., Atal, B. S., Crochiere, R. E., Jayant, N. S., Schroeder, M. R., Tribolet, J. M., "Speech Coding — Digital Techniques for Voice Transmission," IEEE Transactions on Communications, Volume COM-27, April 1979, p. 710-737, Bell Telephone Laboratories, Inc., Murray Hill, New Jersey, Instituto Superior Tecnico, Lisbon, Portugal, Goettingen, Universitaet, Goettingen, West Germany, April 1979.

Flanagan, J. L., Christensen, S. W., "Computer Studies on Parametric Coding of Speech Spectra," Acoustical Society of America, Journal, Volume 68, August 1980, p. 420-430, Bell Telephone Laboratories, Inc., Murray Hill, New Jersey, August 1980.

Flanagan, J. L., "Parametric Coding of Speech Spectra," Acoustical Society of America, Journal, Volume 68, August 1980, p. 412-419, Bell Telephone Laboratories, Inc., Murray Hill, New Jersey, August 1980.

Frehse, R., Geppert, R. Gierloff, R., Kuhn, M. H., Ney H., Piptrowski, H. Tomaschewski, H., "Automatic Recognition of Cooperative Speakers," BM-FT-DV80-009 ISSN-0170-9011, Philips Zentraillaboratorium G.m.b.H., Hamburg, West Germany, December 1980.

Fronczak, E. J., "The ADP Voice Response Interface System and Speech Synthesis Facility," 3rd USA-Japan Computer Conference Proceedings, 231-5, San Francisco, California, October 10-12, 1978, Publ: AFIPS Press, Montvale, New Jersey.

Gabison, A., Gendrin, R., "An Analog Real-Time Analyser for Measuring Instantaneous Amplitude, Frequency and Phase of Signals Varying in Time," Annales des Telecommunications, Volume 34, March-April 1979, p. 158-165 (in French), Office de la Recherche Scientifique et Technique d'Outre-Mer, Paris, France, April 1979.

Georgiou, V. J., "Approaches to Low-Cost Voice Input for Home Computers," IEEE Comput. Soc. Int. Conf. 18th, COMPCON 79 Spring, San Francisco, California, Feb 26-Mar 1, 1979, Publ: IEEE (Cat. No. 79 CH1393-8C), New York.

Geppert, R., Kuhn, M. H., Piotrowsky, H., Tomaschewski, H., "An Operational System for Personnel Authentication Using Long-Term Averaged Voice Spectrum," IEEE Proceedings of the 2979 Carnahan Conference on Crime Countermeasures, 77-80, Lexington, Kentucky, May 16-18, 1979, Publ: Univ. Kentucky, Lexington, Kentucky.

Gettelman, K., "Speaking Up to NC," Mod. Mach. Shop (USA), Volume 53, No. 2, 110-15, July 1980.

Glenn, James W., "Template Estimation for Word Recognition," Proc. IEEE Comput. Soc. Conference Pattern Recognition Image Process, Chicago, Illinois, May 31-June 2, 1978, Publ: IEEE (Cat. No. 78CH1318-5C), New York.

Gold, B., Stern, E., "Information Processing Techniques Program. Volume 1: Packet Speech Systems Technology," Semiannual Technical Summary Report, 1 Apr. - 30 Sep. 1978, AD-A066249 ESD-TR-78-289- VOL-1, Lincoln Lab., Mass. Inst. of Tech., Lexington, September 1978.

Gold, B., Raffel, J. I., Bially, T., "A Study of Future Directions for Low Rate Speech Processor Research, Development, and Implementation," AD-A065063 TN-1978-43 ESD-TR-78-284, Lincoln Lab., Mass. Inst. of Tech., Lexington, November 15, 1978.

Gold, B., "Variable Rate Speech Processing," Digital Signal Processing, London and New York, Academic Press, 1980, MIT, Lexington, Massachusetts, 1980.

Golderos, A., Martinez, R., Nombela, J. R., Pardo, M., Santos, J., Munoz, E., "Man-Machine Voice Communication. I. Speech Production Model," Mundo Electron. (Spain), 47-50, in Spanish, May 1980.

Golderos, A., Martinez, R., Nombela, J. R., Pardo, M., Santos, J., Munoz, E., "Man-Machine Voice Communication. IV. Speech Recognition," Mundo Electron. (Spain), No. 99, 131-4, Coden: MUELCN, September 1980.

Gould, R. L., "Fidelity's Automated Voice Response System," Telecommunications (USA), Volume 15, No. 1, 27-8, Fidelity Service Co., Boston, Massachusetts, January 1981.

Grady, M. W., Hicklin, M. B., Porter, J. E., "Practical Applications of Interactive Voice Technologies: Some Accomplishments and Prospects," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Logicon, Inc., San Diego, California, 1977.

Gray, R. M., Buzo, A., Gray, A. H., Jr., Matsuyama, Y. "Distortion Measures for Speech Processing," IEEE Transactions on Acoustics, Speech, and Signal Processing, Volume ASSP-28, August 1980, Stanford University, Stanford, California, Universidad Nacional Autonoma de Mexico, Mexico City, Mexico, Signal Technology, Inc., Santa Barbara, California, Ibaraki University, Mito, Japan, August 1980.

Greinier, Y., "Speaker Adaptation Through Canonical Correlation Analysis," Dept. Systemes and Communications, Ecole Nat. Superieure Des Telecommunications, Paris, France, IEEE, ICASSP 80 Proceedings, IEEE International Conference on Acoustics, Speech and Signal Processing, 888-91, Part III, Denver, Colorado, April 9-11, 1980, IEEE, New York.

Grooms, D. W., "Applications of the Fast Fourier Transform, Volume 1. A Bibliography with Abstracts," Progress Report, 1970 - Aug 1977, NTIS/PS-78/0931/2, National Technical Information Service, Springfield, Virginia, September 1978.

Gutinger, P., Oberfeuchtner, G., "Rapid and Secure Computer Input," Data Rep. (Germany), Volume 15, No. 5, 19-23, in German, October 1980.

Hadden, D. R., Haratz, D., "The Army Word Recognition System," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Army Communications Research and Development Command, Fort Monmouth, New Jersey, 1977.

Hall, G. T., "Computer Modeling of Voice Signals with Adjustable Pitch and Formant Frequencies," M.S. Thesis, AD-A067556, Naval Postgraduate School, Monterey, California, December 1978.

Harry, D. P., Porter, J. E., Satzer, W. J., "Voice Interactive Analysis System Study," NAVTRAEQUIPC-78-C-0141-1, AD-A074 833/5, NTIS, Logicon, Inc., San Diego, California, Tactical and Training Systems Div. Naval Training Equipment Center, Orlando, Florida, June 1979.

Hartman, W. J., Pratt, L. E., "Voice Performance Measurements," FAA-RD-80-71, National Telecommunications/Information Administration, Boulder, Colorado, Inst. for Telecommunication Sciences, FAA, Washington, D. C., Systems Research and Development Service, June 1980.

Hatfield, J. J., "Cockpit Avionics: Generic," Research and Technology Objectives and Plans Summary; Fiscal Year 1979 Research and Technology Program, NASA-TM-80035, p. 10, (W79-70057, 505-07-23),National Aeronautics and Space Administration, Langley Research Center, Hampton, Virginia, 1979.

Haton, M. C., Haton, J. P., "Sirene, a System for Speech Training of Deaf People," ICASSP 79, 1979 IEEE International Conference on Acoustics, Speech and Signal Processing, 482-5, Washington, D. C., April 2-4, 1979, Publ: IEEE, New York.

Haque, Y. A., "CMOS Operational Amplifiers and Comparator for Use in a Voice Codec with Filters," Asilomar Conference on Circuits, Systems, and Computers (13th) Pacific Grove, California, November 5-7, 1979, Institute of Electrical and Electronics Engineers, American Microsystems, Inc., Santa Clara, California, 1980.

Herscher, M. B., "Real-Time Interactive Speech Technology at Threshold Technology, Inc.," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Threshold Technology, Inc., Delran, New Jersey, 1977.

Herscher, M. B., Simmons, E. J., Welch, J. R., "Voice Data Entry in Medical Applications," Proc. Hawaii Int. Conf. Syst. Sci. 13th, Volume 3, p. 486-501, Honolulu, Hawaii, January 3-4, 1980, Publ: West Period Co., North Hollywood, California.

Hoefker, U., Jesorsky, P., Kriener, B., Talmi, M. Wesseling, D., "New System for Authentication of Voice," Carnahan Conference on Crime Countermeasures Proceedings, Ky. Univ. Off. Res. Eng. Serv. Bull. UKY BU117, p. 47-52, Univ. of Kentucky, Lexington, Ky., May 16-18, 1979, Publ: Univ. of Ky. Coll. of Eng., ORES Publ.

Hoffman, M. Rettberg, R. T., "Development of a Voice Funnel System," AD-A072423 BBN-4143 QTR-2, Bolt, Beranek, and Newman, Inc., Cambridge, Massachusetts, June 14, 1979.

Hogan, J. T., Rozsypal, A. J., "Evaluation of Vowel Duration as a Cue for the Voicing Distinction in the Following Word-Final Consonant," J. Acoust. Soc. Am. (USA), Volume 67, No. 1764-7, Department of Linguistics, Univ. of Alberta, Edmonton, Alberta, Canada, May 1980.

Ibba, G., Paoloni, A., "Description of Main Methodologies Used for Speaker Recognition," FUB-1-1978, in Italian, Fondazione Ugo Bordone, Rome Italy, April, 1976.

Ibba, G., Paoloni, A., "Description of Main Methodologies Used for Speaker Recognition"; "Descrizione Delle Principali Metodologie Utilizzate Per Il Riconoscimento Del Parlatore," FUB-1-1978, NTISNASAE, NTISFNIT, N79-26274/7ST NTIS, in Italian, Fondazione Ugo, Bordoni, Italy, April 1976.

Ibba, G., Paoloni, A., "Automatic Speaker Recognition: Considerations on an Adequate Methodology to Solve the Problem Riconoscimento Automatico Del Parlatore"; "Considerazioni Su Una Metodologica Atta Alla Soluzione Del Problema," FUB-10-1978, Fondazione Ugo Bordoni, Rome, Italy, July 1978.

Ibba, G., Paoloni, A., "Automatic Speaker Recognition: Considerations on an Adequate Methodology to Solve the Problem," FUB-10-1978, in Italian, Fondazione Ugo Bordone, Rome Italy, July 1978.

Ibba, G. Paoloni, A., "Subjective Speaker Recognition: Statistical Consideration on Attainable Results Significance Reconoscimento Soggettivo Del Parlatore"; "Considerazioni Statisiche Sulla Significativita Del Resultati Ottenibili," FUB-17-1978, NTISNASAE, NTISFNIT, N79-26287/9ST, NTIS, in Italian, Fondazione Ugo Bordone, Rome Italy, August 1978.

Ibba, G., Paoloni, A., " Subjective Speaker Recognition: Statistical consideration on Attainable Results Significance," FUB-17-1978, in Italian, Fondazione Ugo Bordoni, Rome, Italy, August 1978.

Institute of Electrical and Electronics Engineers, Inc., Conference on Decision and Control, and Symposium on Adaptive Processes, 16th, and Special Symposium on Fuzzy Set Theory and Applications, Proceedings, Volumes 1 and 2, 1977, Piscataway, New Jersey, 1977.

Input Output Computer Services, Inc., "Single-Channel Voice-Response- System Program Documentation. Volume 3: User's Guide," Final Report, Sept. 1975 - Jan. 1976, AD-A051157, FAA-RD-77-177-3-VOL-3, FAA-77-243-VOL-3, Cambridge, Massachusetts, December 1977.

Input Output Computer Services, Inc., "Single-Channel Voice Response System Program Documentation. Volume 1: System Description," Final Report, Sept. 1975 - Jan. 1976, AD-A051543 FAA-RD-77-177-1VOL-1, FAA-77-24-1-VOL-1, Cambridge, Massachusetts, December 1977.

Input Output Computer Services, Inc., "Single-Channel Voice-Response System Program Documentation. Volume 2: Program-Design Modules," Final Report, Sept. 1975 - Jan. 1976, AD-A051745 FAA-RD-77-177 FAA-77-24-2VOL-2, Cambridge, Massachusetts, December 1977.

Jassem, W. Gembiak, D., Dyczkowski, A., "Computer-Aided Recognition of Polish Vowels in Continuous Speech," Arch. Acoust., Volume 4, No. 1, p. 39-55, Polish Academy of Sci., Inst. of Fundam Technol. Res., Poznan, Poland, 1979.

Jelinek, F., "Continuous Speech Recognition for Text Applications," Electronic Text Communication, 262-76, Munich, Germany, June 12-15, 1978, Publ: Springer-Verlag, Berlin, Germany.

Jewett, W. M., Cole, R., "Modulation and Coding Study for the Advanced Narrowband Digital Voice Terminal TLSP: Interim Report," AD-A060367 NRL-MR-3811, Naval Research Lab., Washington, D. C., August 1978.

Johnson, D. H., Weinstein, C. J., "Phrase Recognizer Using Syllable- Based Acoustic Measurements," IEEE Trans. Acoust. Speech Signal Process, Volume 26, No. 5, p. 409-418, Rice University, Houston, Texas, October 1978.

Johannsen,T. M., "Automated Voice Recognition in Computer Aided Cartography," DCAF E004464, NTIS SAP: HC A-05/MF A 01, N81-33521 24-42, Institut fuer Angewandte Geodaesie, Frankfurt am Main, West Germany, 1980.

Jonnalagadda, K., "Syllabic Compounding and Voice Capacity of a Transponder," RCA Review Volume 41, September 1980, RCA Laboratories, Princeton, New Jersey, 1980.

Kasuya, H., "Comparison of Two Methods for Normalizing Vowel Format Frequencies Based on Estimated Vocal Tract Length," J. Accoust, Soc. Jpn. (E), Volume 1, No. 2, 145-6, Faculty of Engng., Utsunomiya Univ., Utsunomiya, Japan, April 1980.

Kato, Y., Tsurauta, S., "Voice Input for CAD Systems," B82036923, C82027324, Central Res. Labs., Nippon Electric Co. Ltd., Kawasaki, Japan, Proceedings of the IFIP WG5.2-5.3 Working Conf. 203-16, Tokyo, Japan, October 2-4, 1980.

Kimberley, B. P., Searle, C. L., "Automatic Discrimination of Fricative Consonants Based on Human Audition," Rec. IEEE Int. Conf. Acoust. Speech Signal Process 4th, ICASSP '79, p. 89-92, Washington, D. C. April 2-4, 1979, Publ: IEEE (79CH1379-7 ASSP), New York.

Kleiner, R. T., "Automatic Speech Recognition," NASA-CR-104051, Bellcomm, Inc., Washington, D. C., February, 1969.

Klugmann, D., "Talking to the Computer. The Principles and Applications of Automatic Speech Recognition," Central Div. of Technol., Volume 47, No. 4, 16-19, Siemens AG, Munich, Germany, July-August 1980.

Through 1980 (Continued)

Kuhn, Michael H., "Speaker Recognition Accounting for Different Voice Conditions by Unsupervised Classification (Cluster Analysis)," IEEE Trans. Syst. Man Cybern, Volume SMC-10 m, p. 54-57, Philips, Hamburg, Germany, January 1, 1980.

Kurematsu, A., "High Efficiency Speech Coding, Speech Recognition and Speech Synthesis," J. Inst. Telev., Eng. Jpn. (Japan), Volume 33, No. 9, 704-11, in Japanese, Res. and Dev. Labs., Kokusai Denshin Co. Ltd., Tokyo, Japan, September 1979.

Lanier, J. Frank, G., Moron, J. P., "Bathymetric Data Reduction Subsystem (BDRS)," RPT. #: AD-A093747, Defense Mapping Agency Hydrographic and Topographic Center, Washington, D. C., December 1980.

Laycock, J., Peckham, J. B., "Improved Piloting Performance Whilst Using Direct Voice Input," Report # RAE-TR-80019, Royal Aircraft Establishment, Farnborough, England, February 1980.

Lea, W. A., "Critical Issues in Airborne Applications of Speech Recognition TLSP: Final Task Report," AD-A084703, Speech Communication Research Lab., Los Angeles, California, 1979.

Lea, W. A., Shoup, J. E., "Review of the ARPA SUR Project and Survey of Current Technology in Speech Understanding," Final Report, 20 Jul. 1977 - 19 Jul. 1978, AD-A066161, Speech Communication Research Lab., Los Angeles, California, January 1979.

Lea, W. A., "Critical Issues in Airborne Applications of Speech Recognition," AD-A084 703/8, NTIS, Speech Communications Research Lab, Los Angeles, California, 1979.

Lea, W. A., Shoup, J. E., "Review of the ARPA SUR Project and Survey of Current Technology in Speech Understanding," N00014-17-C-0570, AD-A066 161/1ST, NTIS, Speech Communications Research Lab, Los Angeles, California, January 19, 1979.

Lea, W. A., "Commercial Speech Recognizers: A Comparison of their Specifications and Performance Capabilities," Technical Report NAVTRAEQUIUPCEN 79-C-0141-4 to Naval Training Equipment Center, Orlando, FL., Speech Communications Research Laboratory, Los Angeles, CA, September 1980.

Leonard, Gary, "Language Recognition Test and Evaluation," RADC-TR-80-83, Texas Instruments, Inc., Dallas Central Research Labs, Rome Air Development Center, Griffiss AFB, New York, March 1980.

Leonardis, S., "The Processing of Modulation Signal for Digital Radio Transmission," Elektrotehniski Vestnik, Volume 46, Aug.-Oct. 1979, P. 258-261, In Slovene, Ljubljana, Univerza, Ljubljana, Yugolslavia, October 1979.

Lerman, Leon, "Traceability Via Voice Data Entry," IEEE Trans. Components Hybrids Manuf. Technol. CHMT-3, No. 3, p. 360-366, Lockheed Missiles & Space Co. Sunnyvale, California, September 1980.

Lesmo, L., Torasso, P., Mezzalama, M., "Software Methodologies for a Synthesis-by-Rule System — of Man Machine Communication by Voice," Alta Frequenza, Volume 49, May-June 1980, p. 196-205, Torino, Universita, and Torino, Politecnico, Turin, Italy, June 1980.

Levinson, S., Flanagan, J. L. Rabiner, L. R., Rosenberg, A. E., "Applications-Oriented Speech Recognizers," 1978 Wescon Technical Papers, Los Angeles, California, September 12-14, 1978, Publ: Western Periodicals Co., North Hollywood, California.

Levinson, S. E., Shipley, K. L., "A Conversational-Mode Airline Information and Reservation System Using Speech Input and Output," Bell Syst. Tech. J. (USA) Volume 59, No. 1, 119-37, January 1980.

Liberman, A. M., Abranson, A. S., Baer, T., Best, C., Carden. G., "Speech Research, A Report on the Status and Progress of Studies on the Nature of Speech, Instrumentation for Its Investigation, and Practical Applications," SR61(1980), AD-A085 320/-0, Haskins Labs, Inc., New Haven, Connecticut, March 31, 1980.

Lim, J. S., Oppenheim, A. V., "Enhancement and Bandwidth Compression of Noisy Speech," AD-A073139 IN-1979-51 ESD-TR-79-159, Lincoln Lab., Mass. Inst. of Tech., Lexington, June 26, 1979.

Lin, Kun Shan, Gouldi, Kathleen M., Frantz, Gene A., Brantingham, George L., "Text-To-Speech Using LPC Allophone Stringing," E1811195316, IEEE Trans. Connsum, Electron., V. CE-27, N. 2, Pt. 2, P. 144-152, Texas Instrument, Inc., IEEE Chicago, Illinois, November 10-11, 1980.

Link, W., "The Recognition of Speech with Electronic Systems," Nachr. Elektron. (Germany), Volume 32, No. 9, 285-8, in German, September 1978.

Link, W., "Speech Recognition Systems," Nachr, Elektron, (Germany), Volume 33, No. 12, 397-8, in German, December 1979.

Link, W., "Spracheingabe-Systeme - Voice Entry Systems," Nachr Elektron, Volume 33, No. 12, p. 397-398, December 1979.

Lombard, D., Bic, J. C., Imbeaux, J. C., "A Frequency Division Multiple Access Digital Data Transmission System Using Digital Speech Interpolation Equipment and a Viterbi Decoder," International Conference on Digital Satellite Communications Proceedings, 4th, Montreal, Canada, October 23-25, Teleglobe Canada, 1979, p. 306-311, CNET, Issy-les-Moulineaux, Hauts-de-Seine, France, 1979.

Lyle, W. D., Jr., Forte, R., "A Useful Property of the Coefficients of a Walsh-Hadamard Transform," IEEE Transactions on Acoustics, Speech, and Signal Processing, Volume ASSP-28, August 1980, Sandia Laboratories, Albuquerque, New Mexico, Bell Telephone Laboratories, Inc., Whippany, New Jersey, August 1980.

Lyons, J. W., "A Theoretical and Practical Design Investigation of the Future Military Cockpit," International Council of the Aeronautical Sciences, Congress, Proceedings, 12th, Munich, West Germany, October 12-17, 1980, American Institute of Aeronautics and Astronautics, Inc., 1980, British York.

Mariani, J. J., Lienard, J. S., Renard, G., "Speech Recognition in the Context of Two-Way Immediate Person-Machine Interaction," ICASSP 79, 1979 IEEE International Conference on ACOUSTICS, Speech and Signal Processing, 269-72, Washington, D.C., April 2-4, 1979, Publ: IEEE, New York.

Through 1980 (Continued)

Maruta, R., Tomozawa, A., Nishitani, T., "Design of a DSI Terminal for Domestic Applications — Digital Speech Interpolation for Satellite and Terrestrial Communications Systems," NTC '78; National Telecommunications Conference, Birmingham, Alabama, December 3-6, 1978, Conference Record, Volume 1, Institute of Electrical and Electronics Engineers, Inc., 1978, p. 14.3.1-14.3.5, Nippon Electric Co., Ltd., Kawasaki, Japan, 1978.

Matthewson, D. K., "Talking TV (Speech Recognition/Synthesis," Television (GB), Volume 31, No. 2, 74-5, December 1980.

McDowell, Leon, "Source Data Automation Techniques Study for AFMPC," AD-B043 952/1, Report No. RADC-TR-79-317, Rome Air Development Center, Griffiss AFB, New York, December 1979.

McGee, J. P., "Subsystem Status Monitor — for Military Helicopters," AHS 80-17, United Technologies Corp., Sikorsky Aircraft Div., Stratford, Connecticut, American Helicopter Society Annual Forum, A81-40136 18-01, Washington D.C., May 13-15, 1980.

Meoress, M. F., "Sperry Univac Speech Communications Technology," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Sperry Univac, St. Paul, Minnesota, 1977.

Mertz, A. G., "A Preliminary Investigation of Aural Input/Output Systems for In-Flight Information Retrieval TLSP: M.S. Thesis," AD-A063174, Naval Postgraduate School, Monterey, California, September 1978.

Mertz, A. G., "A Preliminary Investigation of Aural Input/Output Systems for In-Flight Information Retrieval," AD=A063 174/7ST, NTIS, Naval Postgraduate School, Monterey, California, September 1978.

Minoli, Daniel, Nakamine, Wendell, "Mersenne Numbers Rooted on 3 for Number Theoretic Transforms," IEEE Int. Conf. Acoust. Speech Signal Process., ICASSP 80 Proc., Volume 1, p. 243-247, Denver, Colorado, April 9-11, 1980, Publ: IEEE (Cat. No. 80CH1559-4), Piscataway, New Jersey.

Minoli, D., "Satellite On-Board Processing of Packetized Voice," ICC '79; International Conference on Communications, Boston, Mass., Conference Record, Volume 4, June 10-14, 1979, Bell Telephone Laboratories, Inc., Holmdel, New Jersey, 1979.

Mirkovic, M., "A Method for Processing of Digital Signals in Order to Reduce the Signal Bandwidth Necessary for Transmission," Archiv fuer Elektronik and Uebertragungstechnik, Volume 33, March 1979, p. 132-134, Elektronska Industrija, Belgrade, Yugoslavia, March 1979.

Modena, G., "'The Speaking Computer' How It Can Synthesize the Human Voice," TE Int. (Italy), Volume 3, No. 2, 12-17, March-April 1979.

Moshier, S. L., "Statistical Assessment of Speech System Performance," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Dialog Systems, Inc., Belmont, Massachusetts, 1977.

Through 1980 (Continued)

Mountford, S. J., North, R. A., "Voice Entry For Reducing Pilot Workload," Human Factors Society Annual Meeting, Los Angeles, California, A81-22901 09-54, Honeywell Systems and Research Center, Minneapolis, Minnesota, October 13-17, 1980.

Mowafi, O. A., Kelly, W. J., Jr., "Integrated Voice/Data Packet Switching Techniques for Future Military Networks," IEEE Trans. Commun. (USA), Volume COM.28, No. 9, Pt. 1, 1655-62, Systems Div., Computer Sci. Corp., Falls Church, Virginia, September 1980.

Mowafi, O. A., Kelly, W. J., Jr., "Integrated Voice/Data Packet Switching Techniques for Future Military Networks," IEEE Trans. Commun. Volume COM-28, No. 9 pt, Comput Sci Corp., Falls Church, Virginia, September 1, 1980.

Mulley, W. G., "Naval Airborne Man-Machine Information Transfer," Proc. S.I.D., (USA), Volume 19, No. 4, 153-5, 1978.

Nassimbene, E. G., "Speech Input Technique for Data Entry to Word Processing Equipment," IBM Tech. Disclosure Bull. (USA), Volume 22, No. 7, 2891-2, IBM Corp., Armonk, New York, December 1979.

Niimi, Y., Kobayashi, Y., "Voice-Input Programming System Using Basic-Like Language," Rec. IEEE Int. Conf. Acoust. Speech Signal Process 3rd, p. 425-428, Tulsa, Oklahoma, April 10-12, 1978, Publ: IEEE (78CH1285-6 ASSP), New York.

O'Neal, J. B., Jr., "Waveform Encoding of Voiceband Data Signals," IEEE Proceedings, Volume 68, February 1980, p. 232-247, North Carolina State University, Raleigh, North Carolina, February 1980.

Oshika, B. T., "FACP Speech Recognition/Transmission System," RADC-TR-78-193, AD-A060 115/3ST, NTIS, System Development Corp., Santa Monica, California, August 1978.

Parker, L. C., Haidt, J. G., "An Automated Pilot Advisory System," Southeastcon -78; Proceedings of the Southeast Region 3 Conference, Atlanta, Georgia, April 10-12, 1978, Institute of Electrical and Electronics Engineers, Inc., 1978, p. 105-108, NASA, Wallops Flight Center, Range Safety Operation Office, Wallops Island, Virginia, 1978.

Parker, L. C., "Digital Operations," Fiscal Year 1979 Research and Technology Program, NASA-TM-80035, p. 10, (W79-70055, 505-07-18), National Aeronautics and Space Administration, Wallops Flight Center, Wallops, Virginia, 1979.

Parker, L. C., "General Aviation Avionics and Controls Research," Research and Technology Objectives and Plans Summary; National Aeronautics and Space Administration, Wallops Flight Center, Wallops Island, Virginia, 1980.

Pay, B. E., "Analysis of Pilots' Voice Records for Automated Speech Recognition," NPL-DNACS-2/78, Avail. NTIS, National Physical Lab., Teddington, England, August 1978.

Pay, B. E., Evans, C. R., "An Approach to the Automatic Recognition of Speech," NPL-DNACS-7/78, Avail. NTIS, National Physical Lab., Teddington, England, October 1978.

Petersen, J. J., Lee, N., Meyn, C., Regelson, E., Satzer, W., "Vocabulary Specification for Automatic Speech Recognition in Aircraft Cockpits," AD-A073 703/1ST, NTIS, Logicon, Inc., San Diego, California, Tactical and Training Systems Div., August 31, 1979.

Petersen, R. J., Lee, N., Meyn, C., Regelson, E., Satzer, W., "Vocabulary Specification for Automatic Speech Recognition in Aircraft Cockpits," Final Report, Sep. 1978 – Jun. 1979, AD-A073703, Logicon, Inc., San Diego, California, August 31, 1979.

Petri, M., "'Speech Treatment' Recognition and Voice Synthesis," Electron, Inc., France, No. 7, 87-91, in French, December 1, 1980.

Pfeifer, L. L., "New Techniques for Text-Independent Speaker Identification," Rec. IEEE Int. Conf. Acoust. Speech Signal Process 3rd, p. 283-286, Tulsa, Oklahoma, April 10-12, 1978, Publ: IEEE (78CH1285-6 ASSP), New York.

Plummer, R. P., Coler, C. R., "Speech as a Pilot Input Medium," Avail. NTIS, MIT Proc., 13th Ann. Conf. on Manual Control, p. 460-462, Utah University, National Aeronautics and Space Administration, Ames Research Center, Moffett Field, California, 1977.

Ponomarev, E. P., "Suboptimal Adaptive Filtering for a Class of Linear Models," 80A32484, Radioteckhnika i Electronika, Vol. 25 (in Russian), April 1980.

Poock, G. K., "Experiments with Voice Input for Command and Control: Using Voice Input to Operate a Distributed Computer Network," NPS55-80-016, Naval Postgraduate School, Monterey, California, April 1980.

Popelka, B. A., Knerr, C. M., "Team Training Applications of Voice Processing Technology," Final Technical Report, 1 Aug. 1979 - 31 Mar. 1980, RPT#: AD-A085999 Litton Mellonics Systems Development Div., Springfield, Virginia, March 31, 1980.

Popelka, Beverly A., Knerr, C. Mazie, "Team Training Applications of Voice Processing Technology," AD-A085 999/1, Litton Mellonics Systems Development Div., Springfield, Virginia, March 31, 1980.

Porter, J. E., "Listen: A System for Recognizing Connected Speech Over Small, Fixed Vocabularies, in Real Time TLSP, Final Report, Jul. 1977 - Mar. 1978," AD-A056231 NAVTRAEQUIPC-77-C-C096-1, Logicon, Inc., San Diego, California, April, 1978.

Portnoy, W. M., Stotts, L. J., "A Miniaturized Digital Telemetry System for Physiological Data Transmission," NASA-CR-160660, Texas Technological Univ., Lubbock, March 31, 1978.

Press, L., "Heuristics Speechlab (Voice Recognition)," Creative Comput. (USA), Volume 5, No. 7, 30-1, July 1979.

Puckorius, T. D., "Illinois' Talking Computer (Telephone Billing System)," Telephony (USA), Volume 195, No. 7, 30, 32, August 14, 1978.

Puckorius, T. D., "Voice Computer System Charges Credit Calls (Telephony)," Signal (USA), Volume 33, No. 4, 27-30, January 1979.

Rabiner, L. R., Wilpon, J. G., Rosenberg, A. E., "Application of Isolated Word Recognition to a Voice Controlled Repertory Dialer System," IEEE Int. Conf. Acoust. Speech Signal Process, ICASSP 80, Proc., Volume 1, p. 182-185, Denver, Colorado, April 9-11, 1980, Publ: IEEE (Cat. No. 80CH1559-4), Piscataway, New Jersey.

Rabiner, L. R., Schmidt, C. E., "Application of Dynamic Time Warping to Connected Digit Recognition," IEEE Trans. Acoust. Speech Signal Process, Volume ASSP-28, No. 4, Bell Lab, Murray Hill, New Jersey, August 4, 1980.

Rabiner, L. R., Wilpon, J. G., Rosenberg, A. E., "A Voice-Controlled Repertory-Dialer System," Bell Syst. Tech. J. (USA), Volume 59, No. 7, p. 1153-63, September 1980.

Ramamoorthy, V., "Voice/Unvoice Detection Based on a Composite-Gaussian Source Model of Speech," ICASSP 80 Proceedings, IEEE International Conference on Acoustics, Speech and Signal Processing, 57-60, Part I, Denver, Colorado, April 9-11, 1980, IEEE, New York.

Reed, W. E., "Vocoders. Citations From the NTIS Data Base," Progress Report, 1964 - Sep. 1979, PB80-801665, National Technical Information Service, Springfield, Virginia, December 1979.

Reimherr, G. W., "Applications of the Fast Fourier Transform, Volume 2. A Bibliography with Abstracts," Progress Report, Sep. 1977 - Aug. 1978, Avail. NTIS, NTIS/PS-78/0932/0, National Technical Information Service, Springfield, Virginia, September 1978.

Reimherr, G. W., "Speech Recognition by Computer. A Bibliography with Abstracts," Report, 1964 - Jun. 1979, NTIS/PS-79/0743, National Technical Information Service, Springfield, Virginia, July 1979.

Reimherr, G. W., "Speech Recognition by Computer (A Bibliography with Abstracts)," NTIS/PS-79/0743/9ST, National Technical Information Service, Springfield, Virginia, July 1979.

Reimherr, G. W., "Speech Recognition by Computer, 1964-June, 1980 (A Bibliography with Abstracts)," PB80-811920, NTIS, National Technical Information Service, Springfield, Virginia, June 1980.

Robinson, A. L., "Communicating with Computers by Voice," IEEE Trans. Prof. Commun. (USA), Volume PC-22, No. 3, 159-65, September 1979.

Sachs, J., Leifer, L., "Voice Command of a Six-Degree-Of-Freedom Manipulator," Proceedings of the 1979 Joint Automatic Control Conference, Denver, Colorado, June 17-21, 1979, Publ: American Inst. Chem. Engrs., New York.

Sakai, R., "Voice Recognition and Synthesis in Information Systems," C81015255, Inc. Process. Soc. Jpn. (Joho Shori), Japan, Vol. 21, No. 8, P. 820-8, Engng, Faculty, Kyoto University, Kyoto, Japan, 1980.

Saridis, G., Lee, C. S. G., "Computer-Controlled Manipulator with Visual Inputs," Opt. Eng., Volume 18, No. 5, p. 492-495, Purdue Univ., West Lafayette, Indiana, September-October 1979.

Sata, Tohio (Ed.), Warman, Ernest (Ed.), "Man-Machine Communication in CAD/CAM," Proceedings of the IFIP WG 5. 2-5 3 Working Conference (Computer Aided Design/Computer Aided Manufacture), University of Tokyo, Japan, 1980.

Sato, M., Murata, M., Namekawa, T., "Differential Pulse Interval and Width Modulated Code," REPT-1525, Technol, Rept. of the Osaka University, Volume 30, No. 1517-1550, p. 73-79, Osaka University, Japan, March 1980.

Saunders, J., "How Voice Data Entry Achieves a High Standard of Recognition," Comput. Wkly. (GB), Volume 25, No. 631, 14-15, December 14, 1978.

Sciulli, J. A., "System Engineering Considerations in DSI Applications – Digital Speech Interpolation," NTC '78; National Telecommunications Conference, Birmingham, Alabama, December 3-6, 1978, Conference Record, Volume 1, Institute of Electrical and Electronics Engineers, Inc., 1978, p. 14.2.1-14.2.5, Telecommunications Techniques Corp., Rockville, Maryland, 1978.

Scott, P. B., "Word Recognition," RADC-TR-78-209, AD-A061 545/0ST, NTIS, Threshold Technology, Inc., Delran, New Jersey, September 1978.

Scott, P. B., "Word Recognition TLSP: Final Report," AD-A061545 RADC-TR-78-209, Threshold Technology, Inc., Delran, New Jersey, September 1979.

Scott, P. B., "DLMS Voice Data Entry TLSP: Final Technical Report, 28 Sep. 1978 - 16 Nov. 1979," AD-A089198 RADC-TR-80-220, Threshold Technology, Inc., Delran, New Jersey, June 1980.

Scott, P. B., "DLMS Voice Data Entry," RADC-TR-80-220, Threshold Technology, Inc., Delran, New Jersey, Rome Air Development Center, Griffiss AFB, New York, June 1980.

Scott, P. B., Welch, J. R., "Low-Cost High-Performance Speech Recognizer for the Handicapped," G2993E4, Fld: 5E, 9B, 95D, GRAI8103, Tribrook Group, Inc. Oak Brook, Ill. May 1980.

Seneff, Stephanie, "Real-Time Harmonic Pitch Detector," IEEE Trans. Acoust. Speech Signal Process, Volume ASSP-26, No. 4, p. 358-365, MIT, Lincoln Lab, Lexington, Massachusetts, August 1978.

Sheshadri, S., Waldron, M. B., "Pattern Recognition Approach to Compare Natural and Synthesized Speech," Rec. IEEE Int. Conf. Acoust. Speech Signal Process 4th, ICASSP '79, p. 777-780, Washington, D. C., April 2-4, 1979, Publ: IEEE (79CH1379-7 ASSP), New York.

Shochet, Ephraim, Delemarre, Louis, Quick, Paul, "Flight Plan Filing by Speech Recognition," AD-A119 679/9, Report No. NADC-79127-60, Naval Air Development Center, Aircraft and Crew Systems Technology Directorate, Warminster, Pennsylvania, January 1979.

Shochet, Ephraim, Lemanski, Richard, "Automated Flight Plan Filing by Simulated Voice Recognition," FAA-RD-80-56, FAA Technical Center, Atlantic City, New Jersey, Aviation Administration, Washington, D. C., Systems Research and Development Service, August 1980.

Shochet, E., Lemanski, R., "Automated Flight Plan filing by Simulated Voice Recognition," Report, Aug. 1975 - May 1979, AD-A089138 FAA-CT-80-15 FAA-RD-80-56, Federal Aviation Agency, Atlantic City, New Jersey, August 1980.

Simmons, E. J., Jr., "Speech Recognition Technology," Computer Design, Volume 18, No. 6, p. 95-101, Threshold Technol. Inc., Delran, New Jersey, June 1979.

Simpson, Carol A., Williams, Douglas H., "Response Time Effects of Alerting Tone and Sematic Context for Synthesized Voice Cockpit Warnings," E1810210450, Human Factors, V. 22, N. 3, P. 319-330, Psycho-Linquist Res. Assoc., Menlo Park, California, June 1980.

Singer Co., "JTIDS Phase 1 Integration Study. Digital Voice Study," AD-B031715L, Little Falls, New Jersey, October 31, 1976.

Skriver, Christian P., "Vocal and Manual Response Modes: Comparison Using a Time-Sharing Paradigm," AD-A119 767/2, Report No. NADC-79127-60, Naval Air Development Center, Aircraft and Crew Systems Technology Directorate, Warminster, Pennsylvania, January 1979.

Sluyter, R. J., Kotmans, H. J., Leeuwaarden, A. V., "A Novel Method for Pitch Extraction from Speech and a Hardware Model Applicable to Vocoder Systems," Philips Res. Labs., Eindhoven, Netherlands, ICASSP 80 Proceedings, IEEE International Conference on Acoustics, Speech and Signal Processing, 45-9, Part I, Denver Colorado, April 9-11, 1980, IEEE, New York.

Smith, C. P., "Intellibility Threshold Level (ITL) Ratings of Some Current Digital Voice Communication Processors," RADC-TR-79-215, Rome Air Development Center, Griffiss AFB, New York, July 1979.

Smith, G. H., "Spectral Peak Amplitude Coding for Voice Data Compression — The Coding of Speech Spectral Bands for Voice Data Processing," Ph.D. Thesis, Univ. Microfilms Order No. 8002017, Stanford Univ., California, 1979.

Sridhar, C. S., Rajappan, K. P., "Simple Pitch Extracter," Z Elektr. Inf. Energietech, Volume 9, No. 1, p. 91-93, Univ. of Cochin, India, 1979.

Steeneken, H. J. M., Agterhuis, E., "Description of Stidas 2-C (Speech Transmission Index Device Using Artificial Signals), Part 1: General Description," IZF-1978-19; TDCK-71457, N79-32450/5 NTIS, Institute for Perception RVO-TNO, Soesterberg, Netherlands, 1978.

Steeneken, H. J. M., Houtgast, T., "A Physical Method for Measuring Speech-Transmission Quality," IZF-1979-1 TDCK-71718, NTIS, Institute for Perception RVO-TNO, Soesterberg, Netherlands, 1979.

Steeneken, H. J. M., Houtgast, T., "A Physical Method for Measuring Speech-Transmission Quality," IZF-1979-1; TDCK-71718, Institute for Perception RVO-TNO, Soesterberg, Netherlands, 1979.

Sugaya, S., Nishimura, S., Smimizu, T., Sugiura, Y., "Voice Remote Controlled TV Set," Compcon 79 Proceedings, Using Microprocessors, Extending our Reach, 392-5, Washington, D. C., September 4-7, 1979, IEEE, New York.

Sykes, D., "Personal Identity Verification," Honeywell Information Systems, Inc., Phoenix, Arizona, Information Privacy (GB, Volume 1, No. 5, 222-5, May 1979.

Takahashi, K., Hishida, E. "The Digital Echo Suppressor for the Experimental Satellite Communication System," IAF Paper 80-D-153, International Astronautical Federation, International Astronautical Congress, 31st, Tokyo, Japan, Nippon Telegraph and Telephone Public Corp., Yokosuka and Tokyo, Japan, September 1980.

Tanaka, H., Chan, C. K., Dressler, M., Dhadesugoor, V. R., Schilling, D. L., Chakravarthy, C. V., "Design of a Packet Voice Transmission System," NTC '79; National Telecommunications Conference, Record Volume 1, Washington D. C., November 27-29, 1979, City College, New York, New York, 1979.

Taylor, H. L., Parikh, H., "A Light-Wand Interface System Design for Quadriplegics," 1980 IEEE Region V Conference Digest, Engineering and Society Interfaces, 115-18, San Antonio, Texas, April 21-23, 1980, IEEE, New York.

Teja, E. R., "Voice Input and Output," EDN (USA), Volume 24, No. 21, 158-67, November 20, 1979.

Tierney, J., "A Study of LPC Analysis of Speech in Additive Noise – Linear Predictive Coding," IEEE Transactions on Acoustics, Speech, and Signal Processing, Volume ASSP-28, August 1980, MIT, Lexington, Massachusetts, August 1980.

Treichler, J. R., "Transient and Convergent Behavior of the Adaptive Line Enhancer," IEEE Transactions on Acoustics, Speech, and Signal Processing, Volume ASSP-27, February 1979, p. 53-62, ARGOSystems, Inc., Sunnyvale, California, February 1979.

Tscheschner, W., "Automatic Speech Processing," Radio Fernsehen Elektron. (Germany), Volume 27, No. 12, 755-9, in German, Tech. Univ. Dresden, Germany, December 1978.

Turmaine, B., "The Principles and Applications of Speech Recognition," Electron. Technol. (GB), Volume 14, No. 10, 224-6, November-December 1980.

Van Der Giet, G., "Applications of the Dawid Speech Recognition System," Sprache and Datenverarb, (Germany), Volume 2, No. 2, 114-21, in German, 1978.

Vasilyev, A. V., Raksheyev, S. S., Ryutina, T. Z., Savelyev, V. P., "System of Vocal Input of Information into an M-222 Computer," FTD- ID(RS)T-218-2180-78, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, December 22, 1978.

Verma, S. N., Ramasastry, J., Monsees, W. R., "Digital Speech Interpolation Applications for Domestic Satellite Communications," NTC '78; National Telecommunications Conference, Birmingham, Alabama, December 3-6, 1978, Conference Record, Volume 1, Institute of Electrical and Electronics Engineers, Inc., 1978, p. 14.4.1-14.4.5, Western Union Telegraph Co., Upper Saddle River, New Jersey, 1978.

Verma, S. N., Maruta, R., Tomozawa, A. Okada, T., Kaul, P., "A Microprocessor Controlled 96/48 Digital Speech Interpolation Terminal," International Conference on Digital Satellite Communications, Proceedings, 4th, Montreal, Canada, October 23-25, 1978, Teleglobe Canada, 1979, p. 261-268, Western Union Telegraph Co., Upper Saddle River, New Jersey, Nippon Electric Co., Ltd., Kawasaki, Japan, Digital Communications Corp., Gaithersburg, Maryland, 1979.

Viglione, S. S., "Multi-User Real Time Word Recognition System," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Interstate Electronics Corp., Anaheim, California, 1977.

Vitols, V. A. "Construction of Adaptive Filters by Use of Subfilters," Asilomar Conference on Circuits, Systems, and Computers (13th) Pacific Grove, California, November 5-7, 1979, Institute of Electrical and Electronics Engineers, Rockwell International Electronics Research Center, Anaheim, California, 1980.

Voiers, W. D., "Toward the Development of Practical Methods of Evaluating Speaker Recognizability," Rec. IEEE Int. Conf. Acoust. Speech Signal Process 4th, ICASSP '79, p. 793-796, Washington, D. C. April 2-4, 1979, Publ: IEEE (79CH1379-7 ASSP), New York.

Wakita, H., "Normalization of Vowels by Vocal-Tract Length and Its Application to Vowel Identification," IEEE Transactions on Acoustics, Speech and Signal Processing, VASSP-25 n2 p183-192, April 1977, Speech Communications Research Lab., Los Angeles, California, April 4, 1976.

Walker, D. E., "Research on Speech Understanding and Related Areas at SRI," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., SRI International Corp., Menlo Park, California, 1977.

Wang J. T., "Design of a Mixed Voice/Data Transmission System for Computer Communication," Ph.D. Thesis, Ohio State Univ., Columbus, 1978.

Wasmeier, M., "Voice Synthesis and Voice Recognition," Nachr. Elektron (Germany), Volume 33, No. 3, 85-6, 88, Datamega KG, Munchen, Germany, in German, March 1979.

Weigel, C. L., "Pilot Preflight Briefing Utilizing an Interactive Computer Generated Voice Response System," Conference on Weather Forecasting and Analysis and Aviation Meteorology, Silver Spring, Maryland, October 16-19, 1978, American Meteorological Society, 1978, p. 324-331, FAA, Washington, D. C., 1978.

Weinstein, C. J., Blankenship, P. E., "Information Processing Techniques Program. Volume 1: Packet Speech Systems Technology," Semiannual Technical Summary Report, 1 Oct. 1979 - 31 Mar. 1980, AD-Aj092113 ESD-TR-80-71 CNT#: F19628-80-C-0002 ARPA, Lincoln Lab., Mass. Inst. of Tech., Lexington, Massachusetts, March 31, 1980.

Welch, J. R., Shamal, E., "Advanced Image Exploitation Aids," RADC-TR-80-74, Threshold Technology, Inc., Delran, New Jersey, Air Development Center, Griffiss AFB, New York, March 1980.

Whelan, J. W., "Digital Voice Processing Overview," NAECON 1979; Proceedings of the National Aerospace and Electronics conference, Volume 3, Dayton Ohio, May 15-17, 1979, Institute of Electrical and Electronics Engineers, Inc., 1979, p. 1304-1311, E-Systems, Inc., St. Petersburg, Florida, 1979.

White, G. M., "Automatic Speech Recognition Technology Development at ITT Defense Communications Division," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Interstate Electronics Corp., Anaheim, California, 1977.

Wilson, S. G., "Adaptive Tree Encoding of Discrete-Time Sources with Speech Applications," NTC '78 National Telecommunications Conference, Birmingham, Alabama, December 3-6, 1978, Conference Record, Volume 2, Institute of Electrical and Electronics Engineers, Inc., 1978, p. 19.5.1-19.5.5, Virginia University, Charlottesville, Virginia, 1978.

Wolf, J. J., "Voice Technology and BBN," Voice Technol. for Interactive Real-Time Command/Control Systems; Naval Air Develop. Center Proc., Bclt, Beranek, and Newman, Inc., Cambridge, Massachusetts, 1977.

Wong, W. C., Steele, R., Goodman, D. J., "Hardware for Detection and Partial Correction of PCM Transmission Errors," IEEE Transactions on Communications, Volume COM-27, Sept. 1979, p. 1310-1315, Loughborough University of Technology, Leics., England, Bell Telephone Laboratories, Inc., Holmdel, New Jersey, September 1979.

Yatagai, A., "Limited and Unlimited Vocabulary Speech Synthesis Systems," E180544629, Telesens Systems, Electron Eng. V. 52, N. 645, P 31-41, London, England, November 1980.

Yoshiya, K. Suzuki, J., "An Equalizer Realized in the Spectral Domain for SPAC/Speech Processing System by Use of Autocorrelation Function," Radio Research Laboratories, Journal, Volume 26, p. 73-81, March 1979.

Youldin, M., Sell, G. H., Reich, T., Clagnaz, M., Louie, H., Kolwicz, R., "Voice Controlled Powered Wheelchair and Environmental Control System for the Severely Disabled," Med. Prog. Technol., Volume 7, No. 2-3, p. 139-143, New York Univ. Med. Center, New York, 1980.

Young, Mary E., "Psycholinguistics (A Bibliography with Abstracts)," PB80-800501, NTIS, National Technical Information Service Springfield, Virginia, November 1979.

Young, Mary E., "Speech Intelligibility, 1964-August, 1980 (Citations from the NTIS Data Base)," PB80-814916, NTIS, National Technical Information Service, Springfield, Virginia, August 1980.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-172590		

4. Title and Subtitle	5. Report Date
STUDY TO DETERMINE POTENTIAL FLIGHT APPLICATIONS AND HUMAN FACTORS DESIGN GUIDELINES FOR VOICE RECOGNITION AND SYNTHESIS SYSTEMS	July, 1985
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
R.W. White and D.L. Parks	D6-52528
9. Performing Organization Name and Address	10. Work Unit No.
Boeing Commercial Airplane Company P.O. Box 3707 Seattle, WA 98124	11. Contract or Grant No. NAS1-17367
	13. Type of Report and Period Covered Contractor Report
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546	14. Sponsoring Agency Code

15. Supplementary Notes

Langley Technical Monitor: George G. Steinmetz
Final Report

16. Abstract

A study was conducted to determine potential commercial aircraft flight deck applications and implementation guidelines for voice recognition and synthesis. At first, a survey of voice recognition and synthesis technology was undertaken to develop a working knowledge base. Then, numerous potential aircraft and simulator flight deck voice applications were identified and each proposed application was rated on a number of criteria in order to achieve an overall payoff rating. The potential voice recognition applications fell into five general categories: programming, interrogation, data entry, switch and mode selection, and continuous/time-critical action control. The ratings of the first three categories showed the most promise of being beneficial to flight deck operations. Possible applications of voice synthesis systems were categorized as automatic or pilot selectable and many were rated as being potentially beneficial. In addition, voice system implementation guidelines and pertinent performance criteria are proposed. Finally, the findings of this study are compared with those made in a recent NASA study of a 1995 transport concept.

17. Key Words (Suggested by Author(s))	18. Distribution Statement		
Voice Recognition Voice Systems Voice Control Digitized Speech Speech Synthesis Phoneme Synthesis Voice Guidelines	Unclassified—Unlimited		
19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	133	