

Opportunities for Advanced Speech Processing in Military Computer-Based Systems*

Clifford J. Weinstein

Lincoln Laboratory, MIT
Lexington, MA 02173-9108

Abstract

This paper presents a study of military applications of advanced speech processing technology which includes three major elements: (1) review and assessment of current efforts in military applications of speech technology; (2) identification of opportunities for future military applications of advanced speech technology; and (3) identification of problem areas where research in speech processing is needed to meet application requirements, and of current research thrusts which appear promising. The relationship of this study to previous assessments of military applications of speech technology is discussed, and substantial recent progress is noted. Current efforts in military applications of speech technology which are highlighted include: (1) narrowband (2400 b/s) and very low-rate (50–1200 b/s) secure voice communication; (2) voice/data integration in computer networks; (3) speech recognition in fighter aircraft, military helicopters, battle management, and air traffic control training systems; and (4) noise and interference removal for human listeners. Opportunities for advanced applications are identified by means of descriptions of several generic systems which would be possible with advances in speech technology and in system integration. These generic systems include: (1) integrated multi-rate voice/data communications terminal; (2) interactive speech enhancement system; (3) voice-controlled pilot's associate system; (4) advanced air traffic control training systems; (5) battle management command and control support system with spoken natural language interface; and (6) spoken language translation system. In identifying problem areas and research efforts to meet application requirements, it is observed that some of the most promising research involves the integration of speech algorithm techniques including speech coding, speech recognition, and speaker recognition.

1 Introduction and Summary

This paper is the result of a study of military applications of advanced speech processing technology which has been undertaken with the following goals: (1) to

review and assess a representative sampling of current efforts in military applications of advanced speech processing technology; (2) to identify opportunities for new military applications, or further development of current applications; and (3) to identify areas where improvements to speech processing technology are needed to address military problems. The intention is to outline a fairly broad range of applications, opportunities and speech technology areas; however, the coverage is not intended to be fully comprehensive, nor to be very detailed in any particular area. Numerous references are provided, on other survey articles, on specific applications and technologies, and on pertinent research efforts.

An historical point of reference for this paper is the 1977 paper by Beek, Neuburg, and Hodge [10], which assessed speech recognition and speech processing technology for military applications. A great deal has been accomplished since 1977 both in speech technology and in applications, but the generic application areas identified in that earlier paper generally remain as relevant now as they were in 1977.

The organization of this paper is as follows. A framework for the paper is first developed by outlining key speech technology areas and key categories of military applications, and identifying which technologies are needed for each category of application. Several previous assessments of military applications of speech technology are then reviewed, and the relationship of this study to those assessments is discussed. A representative sampling of current work in the key military application areas is then summarized and assessed. A set of opportunities for advanced applications are then identified and described; the approach is to describe several generic systems which would be possible with advances in speech technology and in system integration. This description of application opportunities is followed by a brief outline of areas where improvements are needed in speech technology to meet the challenges of military applications, and notes several major technology development efforts which are in progress. Finally, conclusions and areas for further study are discussed.

*This work was sponsored by the Department of the Air Force and the Defense Advanced Research Projects Agency.

2 Framework of Speech Technologies and Military Application Areas

An outline of speech technology areas which are of importance for military (and non-military) applications is presented in Table 1. All these areas are subjects for ongoing research and development. Summaries of the technology are presented in a variety of textbooks and summary papers, and ongoing developments are presented at the annual ICASSP conferences and in other forums. Although the terms used in Table 1 are generally well known, they will be defined, as needed, later in this paper in the context of discussions of particular applications.

Table 1: Speech Technology Areas for Military Applications

1. Speech Recognition
 - 1.1 Isolated Word Recognition (IWR)
 - 1.2 Continuous Speech Recognition (CSR)
 - 1.3 Key-Word Recognition (KWR)
 - 1.4 Speech Understanding (SU)
2. Speaker Recognition
 - 2.1 Speaker Verification (SV)
 - 2.2 Speaker Identification (SI)
 - 2.3 Language Identification (LI)
3. Speech Coding and Digitization
 - 3.1 Waveform Coding
 - 3.2 Source Coding Using Analysis/Synthesis
 - 3.3 Vector Quantization (VQ)
 - 3.4 Multiplexing
4. Speech Enhancement
 - 4.1 Noise Reduction
 - 4.2 Interference Reduction
 - 4.3 Speech Transformations (Rate and Pitch)
 - 4.4 Distortion Compensation
5. Speech Synthesis
 - 5.1 Synthesis from Coded Speech
 - 5.2 Synthesis from Text

Table 2 outlines a number of key military speech application areas which will be addressed in more detail in this paper, and identifies the speech technology areas which are utilized for each application area. The areas in Table 2 generally divide into applications involving speech communication between people, and applications involv-

ing communication between people and computers. In the latter area, speech recognition and synthesis generally would serve as a part of a larger system designed to provide a natural user interface between a person and a computer. Tables 1 and 2 together serve as a framework for the remainder of this paper.

Table 2: Military Speech Application Areas

1. Speech Communications (Speech Coding, Speech Enhancement)
 - 1.1 Secure Communications
 - 1.2 Bandwidth Reduction
2. Speech Recognition Systems for Command and Control (C²) (IWR, CSR, KWR, SU, Synthesis)
 - 2.1 Avionics
 - 2.2 Battle Management
 - 2.3 Resource and Data Base Management
 - 2.4 Interface to Computer and Communication Systems
3. Speech Recognition Systems for Training (IWR, CSR, SU, Synthesis)
4. Processing of Degraded Speech (Enhancement)
5. Security Access Control (SV)

3 Relationship to Previous Assessments of Military Applications of Speech Technology

3.1 Beek, Neuburg, and Hodge (1977)

This paper [10] provided a comprehensive review of the state-of-the-art of speech technology and military applications as of 1977. It serves as a useful reference point for the present paper. The authors grouped potential military applications into four major categories: (1) security (including access control and surveillance); (2) command and control; (3) data transmission and communication; and (4) processing distorted speech. Other than the training application (which was mentioned briefly in the Beek, et al., paper), these categories cover all the application areas listed in Table 2.

Most of the applications cited in the 1977 review were in the research and development stage, and a daunting list of unsolved problems was cited. Much progress has been made since 1977 in speech technology (both algorithms, and hardware implementations of these algorithms) and in applications. The following areas of progress are worthy of particular note:

1. Digital Narrowband Communication Sys-

tems — The Linear Predictive Coding (LPC) algorithm was relatively new in 1977. Improvements in technology and the coding algorithm have now led to widespread deployment of digital narrowband secure voice, especially by means of the STU-III (secure terminal unit) family of equipment at 2.4 kb/s. In addition, significant progress has been made in developing practical coders for lower rates using Vector Quantization (i.e., pattern matching) techniques.

2. **Automatic Speech Recognition** — Major advances both in CSR and IWR have been made largely through the widescale development of statistically-based Hidden Markov Model (HMM) techniques, as well as through the development and application of dynamic time warping (DTW) recognition techniques. HMM techniques which were pioneered prior to 1977, have in recent years been further developed at a large number of laboratories, with significant advances both in recognition performance and in efficiency of implementation. A sampling of basic references on DTW and HMM is provided by [6,21,59,60,113]. [94] provides a good overview of speech recognition technology, and has many useful references. A comprehensive bibliography on speech recognition has recently been published [53].
3. **Noise and Interference Reduction** — Work in application of digital speech processing to noise and interference reduction was relatively new in 1977, and has progressed significantly since that time [89]. Hardware systems for speech enhancement have been developed [28,153] and have been shown to improve both speech readability and ASR performance under certain conditions of noise and interference.

3.2 Woodard and Cupples (1983)

This paper [159] did not attempt a comprehensive review of the state-of-the-art, but instead described selected military applications in three areas: (1) voice input for command and control; (2) message sorting by voice; and (3) very-low-bit-rate voice communications. Current and future applications in the first and third areas will be discussed in some detail in the following discussions. For a general discussion of message sorting and surveillance, the reader is referred to [159].

3.3 Other Reviews of Military Applications of Speech Technology

The 1984 National Research Council report by Flanagan, et al., [38] contains an excellent review of speech recognition system applications to data base management, command and control of weapons systems, and training; a categorization of applications is included, as well as a number of specific case studies. Beek and

Vonusa (1983) [14] provide a general review of military applications of speech technology, with substantial updates from the 1977 Beek, et al., paper referred to above. An early, but comprehensive, assessment of potential military applications of speech understanding systems, is provided by Turn, et al., in 1974 [143]. The book by Lea [68] contains useful discussions on both military and non-military applications of speech recognition. Other applications overviews are presented in [11,12,13]. Taylor (1986) [140] provides a more updated review of avionics applications of speech technology. The Proceedings of Military Speech Technology Conferences (1986–1989) contain a substantial number of useful summaries of specific work in a variety of applications areas. A recent update on military applications of audio processing and speech recognition is provided in [29].

3.4 The NATO Research Study Group on Speech Processing

The North Atlantic Treaty Organization (NATO) Research Study Group on Speech Processing (RSG10) [87], originally formed in 1977, has as one of its major continuing objectives the identification and analysis of potential military applications of advanced technology. In fact the preparation of this paper has been motivated by the author's association with RSG10 since 1986 as a "technical specialist" in the speech area; much useful information for the paper has been provided by other RSG10 members, or learned during RSG10 site visits to various laboratories in the member NATO countries.

The RSG10 group has frequently been involved in the past in activities aimed at disseminating information about speech technology, and military applications in particular, to a wider community. For example, in 1983 the group participated in a NATO Advisory Group for Aerospace Research and Development (AGARD) lecture series on speech processing [2] which included a number of important papers on military applications of speech recognition [14,20]. A similar lecture series was conducted in 1990, and the papers in that series [3] represent an up-to-date overview of a number of important topics in speech analysis/synthesis and recognition systems for military applications.

In another project, RSG10 established in 1983 a working group to look at the human factors aspects of voice input/output systems, which are clearly critical to military (or non-military) applications. A workshop on this subject took place in 1986, resulting in a comprehensive book [141] with papers representative of the state-of-the-art in research and applications in the area of multimodal person/machine dialogs including voice.

In addition to its work in assessing speech technology and opportunities for military applications, RSG10 has continued to initiate and conduct a variety of cooperative international projects [87], particularly in the areas of speech recognition in adverse environments, speech data base collection, speech recognition performance assessment, and human factors issues in speech input/output

systems.

4 Current Work in Development of Military Applications of Speech Technology

Introduction and Summary

This section summarizes and assesses a representative sampling of current work in military applications of speech technology in the following areas: (1) narrowband (2400 b/s–4800 b/s) and low-bit-rate (50–1200 b/s) secure digital voice communications; (2) speech recognition systems in fighter aircraft, military helicopters, battle management, and air traffic control training; and (3) noise and interference suppression.

4.1 Digital Narrowband Secure Voice – the STU-III

Beek, et al., noted in 1977 [10] that “a massive effort is underway to develop and implement an all-digital (secure narrowband speech) communication system.” The development and widespread deployment of the STU-III as described by Nagengast [88] has brought this effort to fruition, and probably represents the single most significant operational military application of speech technology. The STU-III represents a marriage of a sophisticated speech algorithm, the Linear Predictive Coding (LPC) technique at 2.4 kb/s, with very large-scale integration (VLSI) digital signal processor (DSP) technology to allow development of a secure terminal which is small enough and low enough in cost to be widely used for secure voice communication over telephone circuits in the United States. It is worth noting that although the STU-III includes recent improvements in the LPC algorithm, the basic algorithm for 2.4 kb/s LPC has not changed significantly over the last ten years. The primary factor which has allowed its widespread application has been progress in VLSI technology.

Although the 2.4 kb/s LPC algorithm in STU-III produces intelligible speech, it is not toll quality and current efforts are focussed on providing improved quality for secure voice, while maintaining the ability to transmit over standard telephone circuits. Modern technology has evolved to the point where 4.8 kb/s is now generally supportable over the dial network. Hence, recent efforts have focussed, with some success (see [66]) on the development of 4.8 kb/s voice coders with higher quality than LPC. Based on this work, the Code-Excited LPC (CELP) technique has been proposed as a standard for 4.8 kb/s secure voice communication [24]. CELP provides a better representation of the excitation signal, at the cost of a higher bit rate, than the traditional pitch and voiced/unvoiced excitation coding used in 2.4 kb/s LPC.

4.2 Narrowband Secure Voice for Tactical Applications

Most applications of narrowband voice coders at 2.4 kb/s (e.g., STU-III) have been in office environments where background acoustic noise and other environmental effects are not major problems. Operational military platforms such as fighter aircraft, helicopters, airborne command posts, and tanks, pose additional challenges since the performance of narrowband algorithms tend to be sensitive to noise and distortion both in talker and listener environments. However, substantial progress has been made in developing the voice algorithm, microphone, and system integration technology for tactical deployment of 2.4 kb/s voice. Examples include the Joint Tactical Information Distribution System (JTIDS) narrowband voice efforts in the U.S. [123,142] and in the U.K. [125], and the development of the Advanced Narrowband Digital Voice Terminal (ANDVT) family of equipment [134] for a variety of environments.

4.3 Low-Bit-Rate (50–1200 b/s) Voice Communications

Significant advances both in speech algorithms and in VLSI technology have greatly enhanced the feasibility of intelligible, practical digital voice communication at low bit rates (i.e., ≤ 1200 b/s). These coders should have important applicability in a variety of strategic and tactical systems, where channel bandwidth may be extremely limited. Developments in four bit rate ranges are of interest, and will be summarized briefly here. First, it has been demonstrated that frame-fill techniques [16,85,98] can be used very effectively to reduce a 2.4 kb/s algorithm to 1.2 kb/s operation, with little loss in speech performance, and little added complexity. Secondly, to enter the 600–800 b/s range, Vector Quantization (VQ) techniques have been successfully developed [78] which use pattern matching to reduce bit rate. The performance of these VQ systems tends to be sensitive to the speaker on which the patterns are trained, and adaptive training techniques [99] have been developed which effectively adapt the codebook of patterns to the speaker in real time. The third bit rate range of interest is 200–400 b/s. Here, segment vocoder [120] and matrix vocoder techniques which use pattern matching over longer intervals (typically, 100 ms) have been developed. Although quality and intelligibility of these systems are marginal, practical real-time implementations are now feasible, and vocoders at this rate may be useful in selected applications where bandwidth is very limited [49]. Finally, to achieve even lower voice bit rates (say, 50 b/s) it would be necessary to use speech recognition and synthesis techniques, with a restriction on vocabulary. These systems may be useful in selected applications [63] such as transmission of stereotyped reports from a forward observer post. Recognition/synthesis techniques may also be useful for two-way communication in situations where bandwidth is very limited in one direction, but where real-time voice is possible (say, at 1200

b/s) in the other direction [39], allowing confirmation of the correctness of the transmissions which use recognition/synthesis.

4.4 Voice/Data Integration in Computer Networks

The widespread development of computer networks using packet switching technology has opened opportunities for a variety of applications of speech technology, including: packet voice communications [46,148] with efficient sharing of network resources for voice and data; advanced intelligent terminals [39,104] with multi-media communications; multi-media conferencing [104]; and voice control of resources and services (such as voice mail) in computer networks [62,106]. Since data communications using packet systems is becoming widely used in military systems, integration of voice and data on these networks provides significant advantages. Applicable technologies are speech coding, speech recognition, speech synthesis, and multiplexing techniques including (see [148]) Time-Assigned Speech Interpolation (TASI), which take advantage of the bursty nature of speech communications.

4.5 Speech Recognition Systems in High-Performance Fighter Aircraft

The pilot in a high-performance military aircraft operates in a heavy workload environment, where hands and eyes are busy and speech recognition could be of significant advantage. For example, the pilot could use a speech recognizer to set a radio frequency or to choose a weapon, without moving his hands or bringing his gaze inside the cockpit. This would allow the pilot to concentrate more effectively on flying the airplane in combat situations. The potential improvement in pilot effectiveness could be extremely significant in critical situations.

In view of the above, substantial efforts have been devoted over recent years to test and evaluate speech recognition in fighter aircraft. Of particular note are the U.S. program in speech recognition for the Advanced Fighter Technology Integration (AFTI)/F-16 aircraft [55,118,119,121,122,154,158], the program in France on installing speech recognition systems on Mirage aircraft [81,136,137,138], and programs in the U.K. dealing with a variety of aircraft platforms [9,27,41,43,75,128,139,156,157]. In these programs, speech recognizers have been operated successfully in fighter aircraft. Applications have included: setting radio frequencies, commanding an autopilot system, setting steerpoint coordinates and weapons release parameters, and controlling flight displays. Generally, only very limited, constrained vocabularies have been used successfully, and a major effort has been devoted to integration of the speech recognizer with the avionics system.

An excellent description of the roles and limitations of speech recognition systems in fighters, from the user's (i.e., the pilot's) point of view has been presented by

AFTI/F-16 pilot Major John Howard [55]. Several points are worth noting here:

1. speech recognition has definite potential for reducing pilot workload, but this potential was not realized consistently;
2. achievement of very high recognition accuracy (say, 95% or more) was the most critical factor for making the speech recognition system useful – with lower recognition rates, pilots would not use the system;
3. more natural vocabulary and grammar, and shorter training times would be useful, but only if very high recognition rates could be maintained.

With respect to the first point above, the most encouraging result was that for some of the pilots (those for which high recognition rate was achieved), noticeable improvements in overall task performance were achieved with speech recognition for air-to-air tracking and for low-level navigation. A key goal (emphasized by the second point above) is to improve the recognition technology to make these improvements more consistent. Recent laboratory research in robust speech recognition for military environments [100,101,114] has produced promising results which, if extendable to the cockpit, should improve the utility of speech recognition in high-performance aircraft. With respect to the development of vocabularies and grammars which will be well matched to the pilot's needs, a study at the U.S. Air Force Wright-Patterson Avionics Laboratory [76,77] obtained a great deal of useful data by having pilots conduct dialogs with a simulated speech recognition-based system, using mission scenarios simulated in the laboratories. Other discussions of human factors and speech recognition requirements in the cockpit are provided in [15,126].

4.6 Speech Recognition Systems in Helicopter Environments

The opportunities for speech recognition systems to improve pilot performance in military helicopters are similar to those in fighter aircraft. In a hands-busy, eyes-busy, heavy workload situation, speech recognition (as well as speech synthesis) could be of significant benefit to the pilot. Of course, the problems of achieving high recognition accuracy under stress and noise pertain strongly to the helicopter environment as well as to the fighter environment. The acoustic noise problem is actually more severe in the helicopter environment, not only because of the high noise levels but also because the helicopter pilot generally does not wear a facemask, which would reduce acoustic noise in the microphone.

Substantial test and evaluation programs have been carried out in recent years in speech recognition systems applications in helicopters, notably by the U.S. Army Avionics Research and Development Activity (AVRADA) [51,105,115,135,155] and by the Royal Aerospace Establishment (RAE) in the UK [42,74,140]. The program in France has included speech recognition

in the Puma helicopter [138]. Results have been encouraging, and voice applications have included: control of communication radios; setting of navigation systems; and control of an automated target handover system (ATHS) which formats and sends air-air and air-ground messages, and has required a great deal of keyboard entry.

As in fighter applications, the overriding issue for voice in helicopters is the impact on pilot effectiveness. Encouraging results are reported for the AVRADA tests, where it was found [135] that pilots were generally able to run a prescribed course faster and more accurately when speech recognition for radio control was provided. However, these results represent only a feasibility demonstration in a test environment. Much remains to be done both in speech recognition [80] and in overall speech recognition technology, in order to consistently achieve performance improvements in operational settings.

4.7 Speech Recognition Systems in Battle Management

Battle management command centers generally require rapid access to and control of large, rapidly changing information databases. Commanders and system operators need to query these databases as conveniently as possible, in an eyes-busy environment where much of the information is presented in display format. Human-machine interaction by voice has the potential to be very useful in these environments. A number of efforts have been undertaken to interface commercially-available isolated-word recognizers into battle management environments. For example Hale [47] describes the use of a limited vocabulary recognizer for voice recognition control of a weapons control workstation in a command and control laboratory. Although the system capability was limited, the users reported that the voice recognition provided potential convenience in avoiding the need to redirect eyes between screen and keyboard. In another feasibility study [107], speech recognition equipment was tested in conjunction with an integrated information display for naval battle management applications. Again, users were very optimistic about the potential of the system, although capabilities were limited. Another limited application of speech recognition in naval battle management is described in [103].

Clearly, battle management applications of speech recognition systems have high potential; but in order to fully realize this potential, a much more natural speech interface (continuous speech, natural grammar) is needed. The current speech understanding programs sponsored by the Defense Advanced Research Projects Agency (DARPA) in the U.S. has focussed on this problem in the context of a naval resource management task. Speech recognition efforts have focussed on a continuous-speech, large-vocabulary database [108] which is designed to be representative of the naval resource management task. Significant advances in the state-of-the-art in CSR have been achieved, and current efforts are

focussed on integrating speech recognition and natural language processing to allow spoken language interaction with a naval resource management system. Much of this work is described in the Proceedings of recent DARPA Speech Recognition and Natural Language Workshops [31,32,33].

4.8 Training of Air Traffic Controllers

Training for military (or civilian) air traffic controllers (ATC) represents an excellent application for speech recognition systems. Many ATC training systems currently require a person to act as a "pseudo-pilot", engaging in a voice dialog with the trainee controller, which simulates the dialog which the controller would have to conduct with pilots in a real ATC situation. Speech recognition and synthesis techniques offer the potential to eliminate the need for a person to act as pseudo-pilot, thus reducing training and support personnel [20,48,50,127]. Air controller tasks are also characterized by highly structured speech as the primary output of the controller, hence reducing the difficulty of the speech recognition task. The U.S. Naval Training Equipment Center has sponsored a number of developments of prototype ATC trainers using speech recognition. An excellent overview of this work is presented in [20], and further discussion of the results is presented in [38]. Generally, the recognition accuracy falls short of providing graceful interaction between the trainee and the system. However, the prototype training systems demonstrated a significant potential for voice interaction in these systems, and in other training applications. The U.S. Navy is currently sponsoring a large-scale effort in ATC training systems [127], where a commercial speech recognition unit [146] is being integrated with a complex training system including displays and scenario creation. Although the recognizer is constrained in vocabulary, one of the goals of the training programs is to teach the controllers to speak in a constrained language, using specific vocabulary specifically designed for the ATC task. Recent research in France on application of speech recognition in ATC training systems, directed at issues both in speech recognition and in application of task-domain grammar constraints, is described in [82,83,84,91]. In addition to the training application, speech recognition has a variety of other potential applications in ATC systems, as described, for example, in [1].

4.9 Removal of Noise from Noise-Degraded Speech Signals

There are a variety of military and non-military applications where removal of noise and interference from speech signals is important, and a significant amount of work continues to be devoted to this area, both in technology development and in applications. A good summary of the field, with reprints of many important papers, is provided in Lim's book [73]. More recently, a 1989 National Research Council study [89] summarizes

the state-of-the-art in noise removal. Application areas identified in the study include: (1) two-way communication by voice; (2) transcription of a single, important recording; and (3) transcription of quantities of recorded material. The focus of the study is on speech processing to aid the human listener. The panel concluded that, although some noise reduction methods appear to improve speech quality in noise, intelligibility improvements had not been demonstrated using closed response tests such as the Diagnostic Rhyme Test (DRT). The committee recommended further research both on noise reduction algorithm development and on new testing procedures to assess not only intelligibility, but also to assess speech quality, fatigue, workload, and mental effort.

In the area of noise removal, a sustained and successful effort has been sponsored by the Rome Air Development Center [28,152,153,159], which has led to the development of a fieldable development model called the Speech Enhancement Unit (SEU). The SEU has been tested under various realistic noise and interference conditions, and improvements in speech readability have been noted, as well as apparent reduction in operator fatigue.

Prior to any attempts at digital processing for noise removal, it is clearly desirable to apply the most effective possible microphone technology to reduce the noise in the input to the digital system. The effectiveness of standard noise-cancelling microphones is discussed briefly in [142]. Multi-microphone techniques for noise reduction have been the subject of much recent work; [147] and [116] present examples of the work in this area.

In combatting noise and interference in speech broadcast and communication systems, it may be necessary and appropriate in certain situations to process the signal before transmission rather than after reception. Recent work in this area, directed at improving listenability and range in a broadcast system, is described in [110,112].

Finally, recent work has also been successfully directed at advanced headphone technology [19,44] to reduce the noise in the ears of a listener in a high-noise environment. This work has important potential application for speech communications in military environments such as fighter cockpits.

4.10 Speaker Recognition and Speaker Verification

Automatic speech processing techniques for identification of people from their voice characteristics have a number of military and non-military applications, which are summarized in [10] and in [35]. These applications include: (1) security, where the task is to verify the identity of an individual (e.g., for control of access to a restricted facility), and where the subject can often be instructed to speak a required phrase (this is referred to as "text-dependent" speaker verification); (2) surveillance of communication channels [10,35], where the task is to identify a speaker from samples of unconstrained text ("text-independent" speaker recognition); and (3)

forensic applications, which can involve either a recognition or verification task, but where control over the available speech sample is often limited, and the potential number of impostors (i.e., for a verification task) may be very large. Among the above applications, security applications will yield the best speaker recognition performance because of the cooperative user and controlled conditions. A case study of a reasonably successful operational speaker verification system, which has been used to control physical access into a corporate computer center, is described in [35], which points out some of the key problems and solutions in making a successful operational system. Current research efforts in speaker recognition are generally being directed toward the more difficult text-independent speaker recognition problem [35,45,117], with a goal of high performance under conditions of noise and channel distortion.

4.11 Evaluation of Speech Processing Systems

Careful assessment of speech communication systems, speech synthesis systems, and speech recognition systems, using standard data bases and quantitative evaluation measures, is clearly essential for making progress in speech technology for military or non-military applications. Much attention has been directed at the assessment problem in recent years, and an extensive discussion is beyond the scope of this paper. However, the reader is referred to [109] for a comprehensive overview of speech quality assessment, and to [132] for a recent review of evaluation efforts in both speech communication and recognition systems.

5 Opportunities for Advanced Military Applications of Speech Technology

5.1 Introduction and Summary

In this section, opportunities for advanced military applications of speech technology are identified by means of descriptions of several generic systems which would be possible with advances in speech technology and in system integration. These generic systems include: (1) integrated multi-rate voice/data communications terminal; (2) interactive speech enhancement system; (3) voice-controlled pilot's associate system; (4) advanced air traffic control training system; (5) battle management command and control support system with spoken natural language interface; and (6) spoken language translation system.

5.2 Integrated Multi-Rate Voice/Data Communications Terminal

Advanced speech processing will play a very important role in meeting the multiple and time-varying commu-

communications needs of military users. For example, a commander in a fixed or mobile command center will require communication over a variety of networks at a variety of conditions of stress on the networks. An integrated, multi-rate voice/data terminal [39] could be developed to support the commander's needs under normal and stressed conditions as follows: (1) under normal conditions, the terminal would provide secure digital voice, low-rate digital video, and graphics; (2) under heavily stressed conditions with network jamming and damage, the terminal would be limited to stylized data messages; (3) under more favorable but degraded network conditions, more interactive communications would be provided, including very-low-rate secure voice using speech recognition and synthesis techniques.

A sketch of the commander's terminal is shown in Figure 1. The potential roles of advanced speech processing include: (1) a variable-rate coder capable of rates from 50–9600 b/s, depending on network conditions (higher rates, with the attendant higher quality, would be used when conditions permit) and connectivity requirements; and (2) use of speech recognition as an alternate to the keyboard for control of the terminal modes and displays, and for selection or composition of data messages to be transmitted.

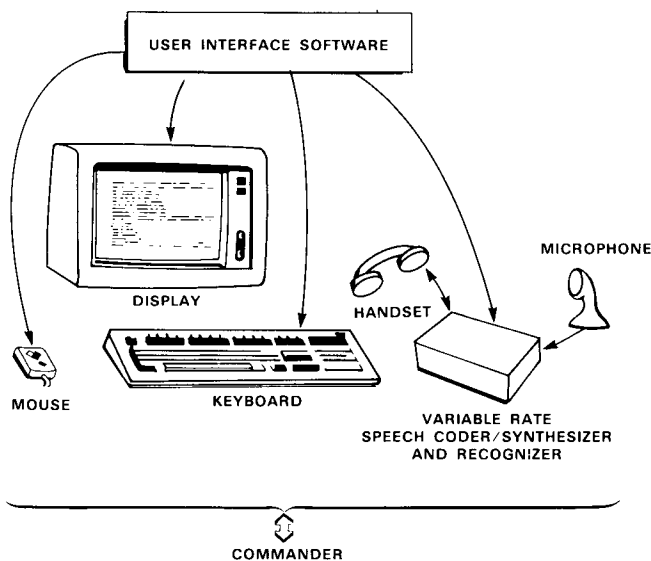


Figure 1: Integrated multi-rate voice/data communications terminal

Variable-rate voice coding, including recognition/synthesis, would also be useful in a scaled-down, very compact terminal for field operations [129] (e.g., by a forward observer in a tactical environment). A sketch indicating this application is shown in Figure 2. The requirements for voice processing are similar to those for the commander's terminal but with a greater emphasis on reduction of size, weight, and power.

The current speech coding technologies discussed above will have to be extended, integrated, and implemented in compact hardware to provide integrated multi-rate terminals for future military communications

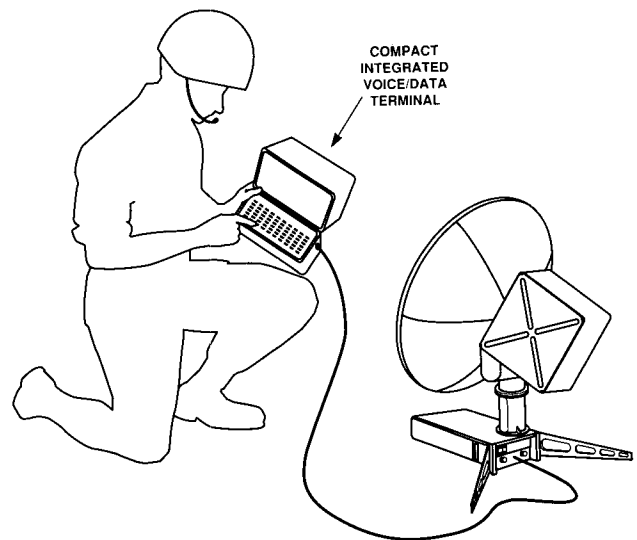


Figure 2: Illustration of portable field terminal concept with variable-rate voice and data communications and data entry/retrieval

needs. In the 200–800 b/s rate range, algorithm and implementation efforts are needed to provide speech coders with good performance. At lower bit rates, improvements to recognition techniques, as well as effective integration of recognition into the communications environment, are needed. As an example, Figure 3 illustrates a concept for recognition-based speech communication in a situation where the two-way link capacities are asymmetric. Here, the outgoing link from one user (e.g., a forward observer with a portable terminal, or an airborne user) might only be able to support rates of 100 b/s or below, while real-time voice (say, at 2400 b/s) is possible in the other direction. The possibility of confirming the recognition/synthesis transmission by means of real-time voice transmission in the reverse direction offers the potential for effective voice communication with disadvantaged links. The development and test of an asymmetric voice coding system of this type could lead to an important military application of advanced speech processing technology.

5.3 Interactive Speech Enhancement Workstation

Advances in speech enhancement technology, coupled with the growing availability of high-performance graphics workstations and signal processing hardware, offer the opportunity for the development of an advanced, interactive speech enhancement workstation with multiple military applications. Such a system, as depicted in Figure 4, would include: (1) real-time speech I/O, including the capability for simultaneous handling of inputs from multiple microphones or sensors [17,37,116,147]; (2) high capacity digital speech storage and playback facilities; (3) a user-selectable library of noise suppression, interference suppression, speech transformation, and filtering

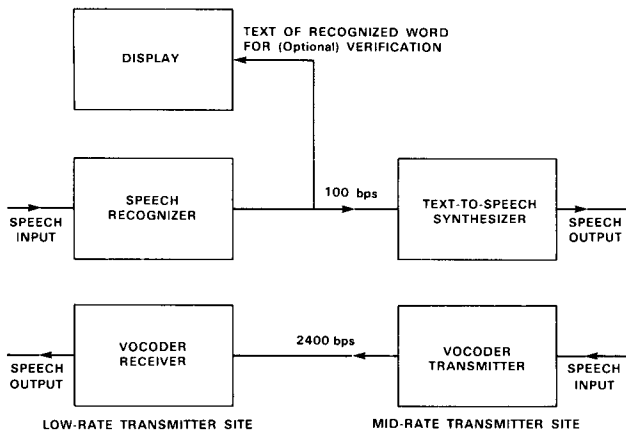


Figure 3: System concept for recognition-based speech communication with asymmetric link capacities

software routines, each capable of operating on real-time speech input or on speech from a digital file; and (4) a user interface providing flexible display, playback, and labelling facilities for speech waveforms, spectra, and parameters.

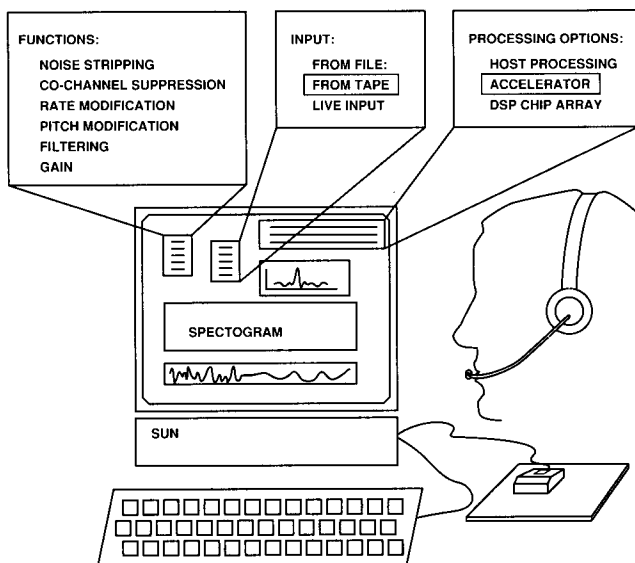


Figure 4: System structure and user interface for interactive speech enhancement workstation

A primary application for such a workstation would be as a listening and transcription aid for degraded speech. As described in [89], two general classes of transcription tasks can be identified: (1) transcription of large quantities of recorded material (such as public broadcasts, or the monitoring of critical telephone lines in a nuclear power station); and (2) transcription of single, important, and often very degraded recording (such as from a cockpit voice recorder after a crash, or forensic

material obtained by a law enforcement agency). The interactive speech enhancement system could be used for either of these transcription tasks, as well as for enhanced listening to real-time speech when transcription is not required.

A great deal of speech enhancement algorithm technology, which would be applicable in such a interactive workstation, has already been developed [28,73,89,152, 153]; and integrating the available algorithms to operate in real-time under flexible user control would be an important development effort.

In addition, there is much to do in the further development of noise and interference reduction, particularly in situations where the interference includes co-channel speech (see, e.g., [90,133,161]).

The advanced interactive speech enhancement workstation represents an important application of advancing speech technology. This work would build on ongoing technology and system efforts, most specifically on the pioneering and ongoing work sponsored by RADC on the speech enhancement unit [28,152], which has included both algorithm development and real-time system implementation using VLSI technology.

5.4 Voice-Controlled Pilot's Associate System

Pilots in combat face an overwhelming quantity of incoming data or communications on which they must base life or death decisions. In addition, they are faced with the need to control dozens of switches, buttons, and knobs to handle the multiple avionics functions in a modern military airplane cockpit. Especially for the case of a single-seater military aircraft, substantial benefit could be achieved through the development of a voice-controlled "pilot's associate", which reduces the pilot's workload, assisting the pilot in controlling avionics system and in keeping track of his changing environment. The concept of the pilot's associate was developed as part of the planning for the DARPA Strategic Computing Program [30], as a paradigm for the development of intelligent "personal associate" systems which could have significant benefits in a variety of human-controlled, complex, military systems.

The pilot's associate would ultimately consist of an ensemble of real-time natural interface system and expert knowledge-based systems. Figure 5 illustrates a concept for an evolving pilot's associate system, which would initially provide a single set of control aids to the pilot, and would evolve to provide a growing set of more complex, knowledge-based functions. In its simplest form, the pilot's associate would include the capability for the pilot to control routine tasks by voice. The efforts described in earlier sections on speech recognition in the cockpit will have to be extended to make speech recognition reliable and useful in the cockpit in order to support functions such as setting radio frequencies, setting navigation systems, or selecting weapons systems.

In its advanced form, the pilot's associate would assist

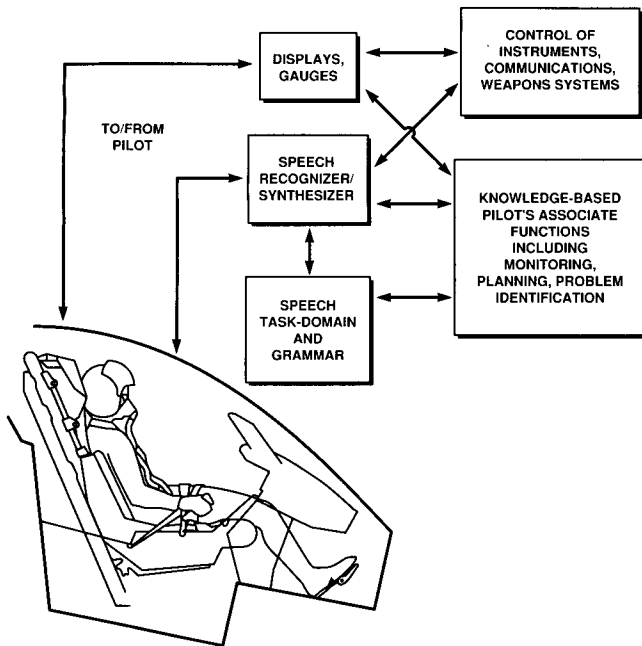


Figure 5: System concept for voice-controlled pilot's associate

the pilot in planning and anticipating functions which would otherwise be very difficult for the pilot, without having a second person on-board. Such functions might include: (1) early detection and diagnosis of an impending malfunction; or (2) presentation of alternate action plans based on the current mission situation. The development of knowledge-based systems to support such tasks presents very difficult challenges, only one of which is an upgrade of the speech recognition interface to improve the naturalness and robustness of pilot interaction with the system.

The pilot's associate represents both an opportunity and a challenge for advanced computing technology in general and for speech technology in particular. The requirement for real-time operation under stressed conditions is particularly demanding for both knowledge-based information monitoring and planning systems, and for the speech interface to the pilot.

5.5 Advanced Air Traffic Control Training System

Automated training systems can use computer speech recognition and generation to expedite training and to reduce the load on training personnel in a variety of applications. Speech recognition and synthesis would be very helpful in hands-busy, eyes-busy training situations, for example in training personnel to maintain complex mechanical equipment. Here the individual could request information from an "automated instruction manual" while continuing to carry on a manual task, and while maintaining his view of the equipment (e.g., a complex jet engine). However, as suggested in Section 4.8, voice-interactive systems are perhaps most attractive for training in tasks which require voice communication as an integral part of the operational task, such as air traffic control (ATC).

Previous efforts in the application of speech technology in ATC training systems have achieved only limited success [20,48,93], but advances in speech technology, simulation technology, expert systems for automated instruction, and performance measurement offer significant potential for major advances in ATC training systems. A generic voice-interactive ATC training system is shown in Figure 6. This particular block diagram was originally drawn to represent a Precision Approach Radar Training System [20], but similar structures would apply to other training scenarios such as air intercept control.

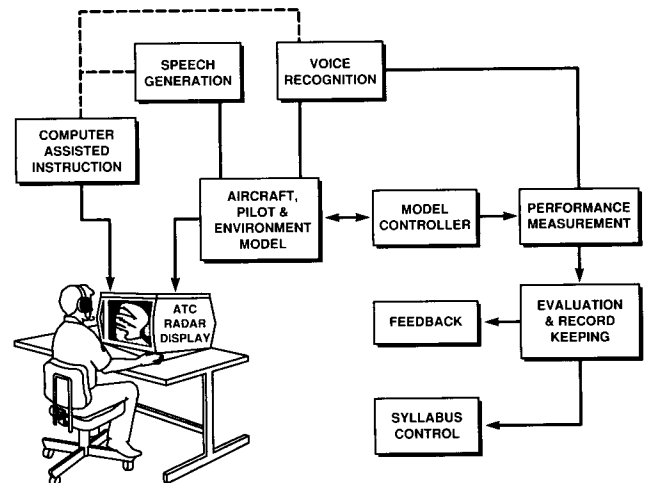


Figure 6: System structure for advanced, automated air traffic control training system with interactive speech recognition/synthesis interface to ATC trainee

The combination of voice-interactive technologies with simulation, environment modelling, and performance measurement has the potential to eliminate the need for a "pseudo-pilot" instructor to interact one-on-one with each student. Automated training has the further advantages of standardizing instruction and of capturing the expertise of the best instructors in the simulated training scenarios. In addition, as new automation capabilities in ATC impose new tasks on the controller (e.g.,[67]), the automated training system could be updated to capture the knowledge of human experts in developing training scenarios which utilize voice-interactive pseudo-pilots.

In the speech technology area, a number of advances will be needed to make an advanced ATC training system effective. Since the controllers are expected to speak in a constrained, stylized language, fully natural speech understanding is not required. However, since controllers will stray from the constraints, it is essential that the recognition system be able to cope effectively with deviations from the constrained vocabulary and grammar. At a minimum, recognition of the deviation and request to the trainee to rephrase his speech input would be needed. Even more desirable would be a system with adaptive training, which learns to extend its vocabu-

lary and grammar based on the trainee's speech to perform correct recognition on an increasing percentage of each trainee's utterances. Adaptive machine learning techniques also offer significant potential in the overall training system, for example in selecting and developing training scenarios which are well-matched to the progress of each ATC trainee.

In summary, the application of speech technology to ATC training is an area of high current interest [18] and significant future potential. In addition, speech recognition and synthesis may have important application in a large variety of intelligent training systems [70], where the computer system effectively simulates a "tutor", communicating with the student in as natural a manner as possible.

5.6 Battle Management Command and Control (C²) Support System with Spoken Natural Language Interface

The application of natural spoken language interfaces in C² systems, including battle management, has been viewed for many years as a long-term goal of speech understanding research including the DARPA speech understanding program in the 1970's [143] and more recent efforts including the DARPA Strategic Computing Program [8,30,33,38,65,96,150]. Some current and previous efforts in this area were noted in an earlier section of this paper. Much remains to be done both in spoken language interface research and in the development of associated support systems and knowledge-based expert systems to support C² users.

Figure 7 shows a sketch of a system for C² battle management with a spoken natural language interface. The generic system structure could be applied to a large variety of C² scenarios [143] including tactical, strategic, and logistics systems; considerable effort over the past few years has been devoted to the application of Naval battle management, under the Fleet Command and Control Battle Management Program (FCCBMP) (see, e.g., [8,34,65,96,150]). There are numerous challenges to be addressed in developing a C² support system with a spoken natural language interface, which include:

1. Techniques for query and management of a large database by spoken natural language must be developed. For the case of FCCBMP, efforts in this area have included: development of the Naval resource management task domain [108], speech understanding work directed at this task domain [31,32,33,34], and porting of natural language interfaces to data base management task for the Naval data base [8].
2. Intelligent expert systems for planning and decision support in the battle management task domain must be developed [65].
3. The spoken natural language interface must be extended to interact with these complex expert systems [65,150].

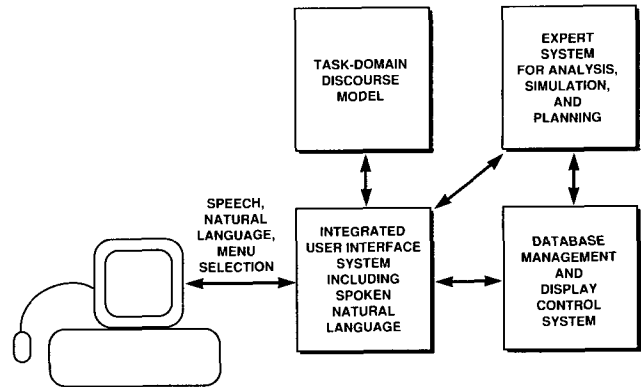


Figure 7: System sketch for C² battle management with a spoken natural language interface. The system includes both relatively simple database retrieval and data entry functions, and more complex expert system aids for battle planning and management. For both classes of functions, the development of a natural spoken language interface represents a considerable challenge, requiring large-vocabulary, natural-grammar speech understanding

4. The speech interface must be combined with other user-interface modalities including graphics, text, and pointing [96].

It is worth emphasizing that although C² systems represent an important opportunity for advanced speech processing in military systems, speech technology development is only one component of the challenge in advanced C² support systems.

Meeting this challenge will require long-term future efforts in speech technology, natural language technology, intelligent system technology, and in system integration. Fortunately, it is not necessary to solve all the problems at once, and a phased approach is possible. For example, initial efforts might involve speech interface to a C² data base management system only (not to the analysis and planning system); the user could initially be required to speak with a constrained vocabulary and grammar while research proceeds on understanding of spoken natural language. Useful aids to commanders and other system users could be provided with the data base management capability only, while work continues on the development and application of the intelligent system technology for the analysis and planning functions needed to provide additional decision aids to the C² user.

5.7 Spoken Language Translation System

Automatic translation of spoken natural language certainly represents one of the "grand challenges" [79] of speech and natural language technology, as well as a long-term opportunity for advanced speech technology. Applications of military relevance include: automatic interpreters for multi-language meetings, NATO field communications, a translating telephone, and translation for cooperative space exploration activities. The impact of automated spoken language translation would clearly be enormous; however, the problem is considerably more difficult than either voice-operated natural language dictation machines or machine translation of text; both of which are unsolved problems requiring much future research. It should be noted, however, that progress continues to be made in dictation systems [7,61]; and new initiatives in machine translation of text are being proposed and developed [54], including application of the powerful statistical techniques [23] which have been successful in speech recognition.

6 Problem Areas for Research

6.1 Introduction and Summary

The Beek, Neuburg, Hodge paper of 1977 [10] concludes with an impressive list of unsolved problems, particularly in the area of automatic speech recognition. The current situation can (perhaps aptly) be summarized by adapting a popular phrase: "You've come a long way, baby, but you've still got a long way to go!". Despite all the progress, much research remains to be done before large-vocabulary continuous speech recognition crosses a threshold of performance sufficient for common use in applications. In other speech technology areas, some real military applications are either at hand or close, but still further research and development efforts are needed to achieve sufficient performance for many other applications.

This section briefly identifies a number of problem areas for research, with a focus on directing attention to references where problems or progress are described in more detail.

A theme sometimes observed in current work, which appears likely to produce significant progress and should be encouraged, is the integration of speech algorithm technologies. For example, speech recognition techniques are applied to speech coding to achieve lower bit rates; and speaker recognition techniques may be integrated with speech coders or speech recognizers to improve robustness of performance across different speakers.

6.2 Low-Rate Speech Coding

Many of the problem areas in low-rate speech coding have already been summarized in earlier sections. At 2.4 kb/s there is a need to move beyond the LPC-10

algorithm toward toll quality speech across a variety of conditions.

At lower rates (i.e., < 800 b/s), improvements in vector quantization [78] and recognition-oriented techniques are needed to make systems effective for general use.

6.3 Noise and Interference Suppression

The state-of-the-art in noise suppression is summarized in [89], which identifies a number of areas for further work in both algorithm development and in evaluation methods. In terms of recent approaches, a variety of combinations of recognition and noise suppression algorithms appear promising [95,36,111]. The suppression of co-channel talker interference is an even more difficult problem than noise suppression [161,90,133], and much work is needed to achieve effective suppression. Following the theme of integration of algorithm technologies, recent work has begun to apply speaker recognition technology to the co-channel interference suppression problem [161,162].

6.4 Speech Recognition in Severe Environments

Prior sections have pointed out both the difficulties and the potential benefits of achieving robust, high-performance speech recognition in severe environments such as fighter aircraft or military helicopters. The National Research Council study [38] report summarizes both the state-of-the-art and research needed for automatic speech recognition in severe environments, as of 1984. Substantial progress has been made since that time, particularly in system development and evaluation on databases of speech collected under stress and noise [114], application of HMM techniques to robust speech recognition [92,100,101,149], and in acoustic-phonetic analysis and compensation for effects of stress and noise [64,130,131]. A number of recent efforts have focussed specifically on compensating for acoustic noise in the HMM recognizer [57,58,144,145]. However, this work has generally been performed for severe conditions which are simulated in the laboratory, and has achieved best performance for isolated-word recognition. Much work remains to achieve high-performance, continuous speech recognition under severe operational conditions; an essential, though costly, requirement for achieving progress in this area is a continuing program of data collection and speech recognizer testing in real (e.g., fighter or helicopter) military environments.

6.5 Large-Vocabulary Continuous Speech Recognition

There has been a great deal of effort and much progress in the area of large-vocabulary continuous speech in recent years [5,31,32,33,34]. But substantial improvements in performance are still needed before such systems achieve high enough accuracy to be usable in practical applications [79]. For example, a February 1989

evaluation [97] of a number of state-of-the-art systems on the 1000-word, perplexity-60 DARPA resource management task yielded the following best results:

1. speaker-dependent: word error rate 3.1%, sentence error rate 21.0%;
2. speaker-independent: word error rate 6.1%, sentence error rate 34.3%.

(Perplexity is a measure of the recognition task difficulty, and is defined as the probabilistically-weighted geometric mean branching factor of the language (see, e.g., [69], pp. 145-146)). For a 5000-word, perplexity-93 task, recent systems have achieved a speaker-dependent word error rate of 11.0% [5]. For an aggressive (but not unrealistic for applications requirements) goal, such as 95% speaker-independent sentence recognition for a 5,000-word vocabulary system, it is clear that an order-of-magnitude improvement in word error rate is needed. Some potential sources of improvement, where research is needed, include [79] better signal representation, better modelling of linguistic units, and better parameter estimation. Recent efforts in phonetic classification using neural networks [71], and in combining neural network pattern classifiers with HMM techniques [22,40,56,72,160] offer potential promise in these areas. Additional system performance improvements should be achieved by improved language modelling, and by the integration of speech recognition and natural language processing systems [34], as discussed further below.

6.6 Natural Language Processing Technology

Advanced systems involving interaction of people with computers using spoken language will clearly require substantial advances in natural language processing. A summary of the state-of-the-art, major challenges, and research opportunities in natural language processing is presented in [151]. A major "grand challenge" cited for natural language processing is support for natural, interactive dialogue; this challenge, which must be addressed even for textual natural language input, is clearly a pressing and more difficult challenge for spoken natural language input.

6.7 Integration of Speech Recognition and Natural Language Processing Systems

The integration of speech recognition and natural language processing systems to form interactive spoken language systems is a key focus for current [34] and future [79,151] research, needed to develop practical systems for complex military (and other) applications. Specific recent efforts which indicate promising directions for research and development in integration of speech recognition and natural language processing are described in [86,102,124,163].

6.8 Human Factors

Work on human factors integration is an often neglected, but crucial, element which is necessary to realize the benefits of speech technology. Major areas of concern, which are summarized in [38], include task selection, dialogue design, user characteristics, speech display design, task environment, and overall system performance assessment. Recent progress and research directions in the specific area of person/machine dialogues are covered rather broadly in [141].

An excellent controlled study in the use of voice dialogs to accomplish specific tasks, which illustrates the effectiveness of voice as compared with other modalities (e.g., typing, writing), is described in [25]. Experiments indicating the capability of people to restrict themselves to limited vocabularies in task-oriented dialogues are described in [26]. In general, more attention is needed to human factors and dialogue issues, and speech system developers can benefit significantly from the results of studies such as these, and from other studies cited in [141].

6.9 Speaker Recognition and Verification

An excellent summary of the state-of-the-art and of applications in speaker recognition and verification, as of 1985, is provided in [35]. Some promising recent efforts in the challenging problem of text-independent speaker recognition are described in [45,117]. An additional important current (and projected future) research thrust [79] is the application of speaker recognition techniques to adapt speech recognizers to speaker characteristics, and hence to improve speech recognition performance.

7 Summary and Conclusions

There has been much progress in recent years in speech technology and in application of this technology in military systems. This progress has brought some applications of speech technology into operational use, has brought other applications into the development and test stage, and has brought closer the opportunities for advanced applications. To realize the necessary technology for advanced applications, and to bring advanced applications into practice, much work is needed in basic speech algorithm technology, speech system implementation, and iterative test and improvement of fielded systems.

This paper has presented in overview form: (1) a review and assessment of current military applications of speech technology; (2) an identification of a sampling of opportunities for future military applications of advanced speech technology; and (3) an identification of problem areas for research to meet applications requirements, and of promising research thrusts.

Applications in narrowband speech communication and speech enhancement are seen to be at hand, and opportunities as described for advanced voice/data work-

stations, based on extension and integration of current technology. In the speech recognition and spoken language understanding areas, a number of current applications are described which are generally in the development and test stage. A number of opportunities for advanced applications in these areas are described; these generally will require significant advances in speech recognition technology, and in the integration of speech recognition into systems which will require advanced natural language processing and careful attention to system integration and human factors. Current efforts which are underway in speech recognition, natural language processing, and system development technology should lead, over the next several years, to significant technology advances in speech technology, and to significant progress toward realizing these and other opportunities for military applications of advanced speech processing.

Acknowledgment

I would like to acknowledge the contributions to this paper of my colleagues on the NATO Research Study Group in Speech Processing (RSG10) in providing technical suggestions, identifying applications and pertinent references, and providing valuable comments and criticism. I would specifically like to acknowledge the contributions and encouragement of the U.S. delegate, Jim Cupples, and of the past and current RSG10 chairmen, Roger Moore and Herman Steeneken.

References

- [1] M.G. Abbott, "The Use of Speech Technology to Enhance the Handling of Electronic Flight Progress Strips in an Air Traffic Control Environment," Proc. Voice Systems Worldwide, London, Media Dimensions, pp. 126-134, May 1990.
- [2] Proceedings of the NATO AGARD Lecture Series No. 129 on "Speech Processing," June 1983.
- [3] Proceedings of the NATO AGARD Lecture Series No. 170 on "Speech Analysis and Synthesis and Man-Machine Speech Communications for Air Operations," May 1990.
- [4] L.R. Bahl, et al., "Large Vocabulary Natural Language Continuous Speech Recognition," IEEE Proc. ICASSP'89, pp. 465-468, May 1989.
- [5] L.R. Bahl, et al., "Large Vocabulary Natural Language Continuous Speech Recognition," in IEEE Proc. Int. Conf. Acoustics, Speech and Signal Processing 1989, pp. 465-468, May 1989.
- [6] J.K. Baker, "The Dragon System — An Overview," IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-23, No. 1, pp. 24-29, February 1975.

- [7] J.M. Baker and J.K. Baker, "The Dragon-Dictate System," demonstration at the DARPA Speech and Natural Language Workshop, see Conf. Proc. (Morgan-Kaufmann), February 1989, p. 11.
- [8] M. Bates, "Rapid Porting of the Parlance (TM) Natural Language Interface," Proc. DARPA Speech and Natural Language Workshop (Morgan-Kaufmann), February 1989, pp. 83-88.
- [9] P. Beckett, "Voice Control of Cockpit Systems," NATO AGARD Conf. Proc. No. 414 Information Management and Decision Making in Advanced Airborne Weapon Systems, Toronto, Canada, April 1986.
- [10] B. Beek, E.P. Neuburg, and D.C. Hodge, "An Assessment of the Technology of Automatic Speech Recognition for Military Applications," IEEE Trans. Acoustics, Speech and Signal Processing, Vol. 25, No. 4, pp. 310-321, August 1977.
- [11] B. Beek and E.J. Cupples, "Military Applications of Automatic Speech Recognition and Future Requirements," Proc. Conf. on Voice Technology for Interactive Real-Time Command/Control Systems Applications, pp. 197-206, NASA-AMES Research Center, CA, December 1977.
- [12] B. Beek, E.J. Cupples, J. Ferrante, J. Nelson, J. Woodard and R. Vonusa, "Trends and Application of Automatic Speech Technology," Proc. Symposium on Voice Interactive Systems Applications and Payoffs, pp. 63-72, Naval Air Development Center, PA, May 1980.
- [13] B. Beek, "Overview of State-of-the-Art, R&D NATO Activities, and Possible Applications — Voice Processing Technology," Proc. AGARD Conf. No. 329 on Advanced Avionics and the Military Aircraft Man/Machine Interface, April 1982.
- [14] B. Beek and R.S. Vonusa, "General Review of Military Applications of Voice Processing," Proc. AGARD Lecture Series No. 129 on Speech Processing, June 1983.
- [15] R. Bell, M.E. Bennett, and W.E. Brown, "Direct Voice Input for the Cockpit Environment," Proc. AGARD Conf. No. 329 on Advanced Avionics and the Military Aircraft Man/Machine Interface, April 1982.
- [16] P.E. Blankenship and M.L. Malpass, "Frame-Fill Techniques for Reducing Vocoder Data Rates," Technical Report 556, Lincoln Laboratory, M.I.T., February 1981.
- [17] S.F. Boll and D.C. Pulsipher, "Suppression of Acoustic Noise in Speech Using Two-Microphone Adaptive Noise Cancellation," IEEE Trans. Acoust., Speech and Signal Processing, ASSP-28, pp. 752-753, 1980.

- [18] G. Booth, Chairman, Session on Training and Air Traffic Control in Proc. Speech Tech 90 (Media Dimensions), April 1990.
- [19] A.G. Bose and J. Carter, "Headphoning," U.S. Patent No. 4,455,675, June 19, 1984.
- [20] R. Breaux, M. Blind, and R. Lynchard, "Voice Technology in Navy Training Systems," Proc. AGARD Lecture Series No. 129 on Speech Processing, Jun 1983.
- [21] J.S. Bridle, R.M. Chamberlain, and M.D. Brown, "An Algorithm for Connected Word Recognition," Proc. ICASSP'82 (Paris).
- [22] J.S. Bridle, "Alpha-Nets: A Recurrent 'Neural' Network Architecture with a Hidden Markov Model Interpretation," Royal Signals and Radar Establishment Research Note SP4, October 1989; also to appear in *Special Neurospeech Issue on Speech Communications*, 1990.
- [23] P. Brown, et al., "A Statistical Approach to French/English Translation," Proc. RIAO 88 Conf. on User-Oriented Content-Based Text and Image Handling, Massachusetts Institute of Technology, March 1988, pp. 810-823.
- [24] J.P. Campbell, V.C. Welch, and T.E. Tremain, "An Expandable Error-Protected 4800 bps CELP Coder (U.S. Federal Standard 4800 bps Voice Coder)," Proc. ICASSP'89, Glasgow, pp. 735-738.
- [25] A. Chapanis, "Interactive Human Communication," *Scientific American*, Vol. 232, pp. 36-49, March 1975.
- [26] A. Chapanis, "Interactive Communication: A Few Research Answers for a Technological Explosion," in D. Neel and J.S. Lienard, eds., *Nouvelles Tendances de la Commucation Homme-Machine/New Trends in Human-Machine Communication*, INRIA, Le Chesnay, 1980.
- [27] N. Cooke, "RAE Bedford's Experience of Using Direct Voice Input (DVI) in the Cockpit," Proc. Voice Systems Worldwide, pp. 135-153, Media Dimensions, London, England, May 1990.
- [28] E.J. Cupples and J.L. Foelker, "Air Force Speech Enhancement Program," Proc. Military Speech Tech 1987 (Media Dimensions) November 1987.
- [29] E.J. Cupples and B. Beek, "Application of Audio/Speech Recognition for Military Applications," Proc. AGARD Lecture Series No. 170 on "Speech Analysis and Synthesis and Man-Machine Speech Communications for Air Operations," May 1990.
- [30] Defense Advanced Research Projects Agency Report, "Strategic Computing — New Generation Computing Technology: A Strategic Plan for Its Development and Application to Critical Problems in Defense," 28 October 1983.
- [31] Proceedings of the DARPA Speech Recognition Workshop, Science Applications International Corp., Report No. SAIC-86/1546, February 1986.
- [32] Proceedings of the DARPA Speech Recognition Workshop, Science Applications International Corp., Report No. SAIC-87/1644, March 1987.
- [33] Proceedings of the February 1989 DARPA Speech and Natural Language Workshop (Morgan Kaufman), February 1989.
- [34] Proc. DARPA Speech and Natural Language Workshop (Morgan-Kaufmann), October 1989.
- [35] G.R. Doddington, "Speaker Recognition — Identifying People by their Voices," Proc. of the IEEE, Vol. 73, No. 11, pp. 1651-1664, November 1985.
- [36] Y. Ephraim, D. Malah, and B.H. Juang, "On the Application of Hidden Markov Models for Enhancing Noisy Speech," IEEE Proc. ICASSP'88, pp. 533-536, April 1988.
- [37] M. Feder, A.V. Oppenheim, and E. Weinstein, "Methods for Noise Cancellation Based on the EM Algorithm," IEEE Proc. ICASSP'87, pp. 201-204, April 1987.
- [38] J.L. Flanagan, N.R. Dixon, G.R. Doddington, J.I. Makhoul, M.E. McCauley, E.F. Roland, J.C. Ruth, C.A. Simpson, B.H. Willeges, W.A. Woods, and V.W. Zue, "Automatic Speech Recognition in Severe Environments," National Research Council Committee on Computerized Speech Recognition Technologies Report (National Academy Press, Washington, D.C.) 1984.
- [39] J.W. Forgie, A.J. McLaughlin, and C.J. Weinstein, "Architecture for an Intelligent C³ Terminal," Proc. MILCOM'87, October 1987.
- [40] M. Franzini, K.F. Lee, and A. Waibel, "Connectionist Viterbi Training: a New Hybrid Method for Continuous Speech Recognition," in IEEE Proc. Int. Conf. Acoustics, Speech and Signal Processing 1990, pp. 425-428, April 1990.
- [41] I. Galletti and M. Abbott, "Advanced Airborne Speech Recogniser," Proc. American Voice I/O Soc. Conf. AVIOS-89, Newport Beach, CA, pp. 127-136, September 1989.
- [42] I. Galletti and M. Abbott, "Advanced Airborne Speech Recognizer," AVIOS Conf. Proc., September 1989.
- [43] I. Galletti and M. Abbott, "Development of an Advanced Airborne Speech Recogniser for Direct Voice Input," *Speech Technology Magazine*, Vol. 5, No. 1, pp. 60-63, October/November 1989.

- [44] D. Gauger and R. Sapiejewski, "The Bose Headset System: Background, Description, Applications," Bose Corporation Technical Report, January 1987.
- [45] H. Gish, "Robust Discrimination in Automatic Speaker Identification," IEEE Proc. ICASSP'90, pp. 289-292, April 1990.
- [46] B. Gold, "Digital Speech Networks," Proc. IEEE, Vol. 65, pp. 1636-1658, November 1977.
- [47] M. Hale and O.D. Norman, "An Application of Voice Recognition to Battle Management," Proc. Military Speech Tech 1987 (Media Dimensions), Arlington, VA, November 1987.
- [48] J.A. Harrison, G.R. Hobbs, J.R. Howes, N. Cope, and G.R. Wright, "Machine Supported Voice Dialogue Used in Training Air Traffic Controllers," Proc. IEE Int. Conf. on Speech Input/Output: Techniques and Applications, No. 258, pp. 110-115, March 1986.
- [49] J.R. Herman, C.C. Duggan, W.I. Thompson, R.A. Costa, and D.M. Deluca, "Narrowband Digital Voice Communications Over a Meteor Burst Channel," Electronics Letters, Vol. 23, No. 1, January 1987.
- [50] G.R. Hobbs, "The Application of Speech Input/Output to Training Simulators," Proc. IFS Conf. on Speech Technology, pp. 121-134, Brighton, UK, October 1984.
- [51] J. Holden, S.J. Stewart, and G. Vensko, "Speech Recognition in a Helicopter Environment," Proc. Military Speech 1987 (Media Dimensions), Arlington, VA, November 1987.
- [52] J.M. Holden, "Testing Voice I/O in Helicopter Cockpits," AVIOS Conf. Proc., September 1989.
- [53] A.S. House, *The Recognition of Speech by Machine - A Bibliography* (Academic Press) 1988.
- [54] E.H. Hovy, "New Possibilities in Machine Translation," Proc. DARPA Speech and Natural Language Workshop (Morgan-Kaufmann), October 1989, pp. 99-112.
- [55] J.A. Howard, "Flight Testing of the AFTI/F-16 Voice Interactive Avionics System," Proc. Military Speech Tech 1987, (Media Dimensions) November 1987.
- [56] W.Y. Huang and R.P. Lippmann, "HMM Speech Recognition with Neural Net Discrimination," in IEEE Proc. Neural Information Processing Systems 1989 (NIPS'89), Boulder, CO, November 1989.
- [57] M.J. Hunt and C. Lefebvre, "A Comparison of Several Acoustic Representations for Speech Recognition with Degraded and Undegraded Speech," Proc. ICASSP'89, Glasgow, pp. 262-265, May 1989.
- [58] M.J. Hunt and S.M. Richardson, "The Use of Linear Discriminant Analysis in a Speech Recognizer," Proc. Voice Systems Worldwide, London, Media Dimensions, pp. 87-93, May 1990.
- [59] F. Itakura, "Minimum Prediction Residual Applied to Speech Recognition," IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-23, pp. 67-72, 1975.
- [60] F. Jelinek, "Continuous Speech Recognition by Statistical Methods," IEEE Proceedings, Vol. 64, No. 4, pp. 532-556, April 1976.
- [61] F. Jelinek, "The Development of an Experimental Discrete Dictation Recognizer," Proc. IEEE, Vol. 73, No. 11, pp. 1616-1624, November 1985.
- [62] D.H. Johnson and C.J. Weinstein, "A User Interface Using Recognition of LPC Speech Transmitted Over the ARPANET," EASCON Conf. Record, September 1977.
- [63] R.R. Johnson and S.W. Nunn, "Voice I/O Applications at the Naval Oceans Systems Center," Proc. Speech Tech '86, pp. 255-257, New York City, April 1986.
- [64] J. Junqua and Y. Anglade, "Acoustical and Perceptual Studies of Lombard Speech: Application to Isolated-Word Automatic Speech Recognition," Proc. ICASSP'90, Albuquerque, pp. 841-844, April 1990.
- [65] S.H. Kaisler, "FRESH: A Command and Control Prototypes," Command, Control, and Communications Technology Assessment: Conf. Proc., Defense Communications Agency, November 1986, pp. III-1-III-8.
- [66] D.P. Kemp, R.A. Sueda, and T.E. Tremain, "An Evaluation of 4800 bps Voice Coders," Proc. ICASSP'89, pp. 200-203 (Glasgow).
- [67] R.R. LaFrey, "Parallel Runway Monitor," *Lincoln Laboratory Journal*, Vol. 2, No. 3, pp. 411-436, Fall 1989.
- [68] W.A. Lea, *Trends in Speech Recognition*, (Prentice Hall, Englewood Cliffs, NJ), 1980. (See Part 1, motivations and general reviews.)
- [69] K.F. Lee, *Automatic Speech Recognition: The Development of the SPHINX System*, Kluwer Academic Publishers, 1989.
- [70] A. Lesgold, S. Chipman, J. Brown, and E. Soloway, "Intelligent Training Systems," in *Annual Review of Computer Science 1990*, Vol. 4, pp. 383-394, Annual Reviews, Inc., 1990.

- [71] H.L. Leung and V.W. Zue, "Phonetic Classification Using Multi-Layer Perceptrons," in IEEE Proc. Int. Conf. Acoustics, Speech and Signal Processing 1990, pp. 525-528, April 1990.
- [72] E. Levin, "Word Recognition Using Hidden Control Network Architecture," in IEEE Proc. Int. Conf. Acoustics, Speech and Signal Processing 1990, pp. 433-436, April 1990.
- [73] J.S. Lim (ed.), *Speech Enhancement* (Prentice-Hall) 1983.
- [74] R. Little and R. Cowan, "A Flight Evaluation of Voice Interaction as a Component of an Integrated Helicopter Avionics System," RAE Technical Memo FS(B) 637, April 1986.
- [75] R. Little, "The Flight Evaluation of a Speech Recognition and a Speech Output System in an Advanced Cockpit Display and Flight Management System for Helicopters," NATO AGARD Conf. Proc. No. 414 on Information Management and Decision Making in Advanced Airborne Weapon Systems, Toronto, Canada, April 1986.
- [76] G. Lizza and C.J. Goulet, "Cockpit Natural Language - An Application Specification," Proc. IEEE National Aerospace and Electronics Conf., NAECON, pp. 818-819, 1986.
- [77] G. Lizza, M. Munger, R. Small, G. Feitshans and S. Detro, "A Cockpit Natural Language Study," AFWAL TR-87-3003, April 1987.
- [78] J. Makhoul, S. Roucos, and H. Gish, "Vector Quantization in Speech Coding," Proc. IEEE, Vol. 73, No. 11, pp. 1551-1588, November 1985.
- [79] J. Makhoul, F. Jelinek, L. Rabiner, C. Weinstein, and V. Zue, "White Paper on Spoken Language Systems," Proc. DARPA Speech and Natural Language Workshop (Morgan-Kaufmann), October 1989, pp. 463-480; also published in Annual Review of Computer Science 1990, Vol. 4, pp. 481-501, Annual Reviews, Inc., 1990.
- [80] F.J. Malkin and T.W. Dennison, "The Effect of Helicopter Vibration on the Accuracy of a Voice Recognition Systems, Proc. IEEE National Aerospace and Electronics Conf., NAECON, pp. 813-817, 1986.
- [81] J. Mariani, B. Prouts, J.L. Gauvain, and J.J. Gangolf, "A Man-Machine Speech Communication System Including Word-based Recognition and Text-to-Speech Synthesis," IFIP 1983, Paris, pp. 673-679, September 1983.
- [82] K. Matrouf, F. Néel, J. Mariani, C. Bailleul, P. Dujardin, and G. Knibbe, "Prototype Industriel de Système de Dialogue dans le Contrôle Aérien/Industrial Prototype for a Vocal Dialogue in Air Traffic Control," AFCET/INRIA, 7ème Congrès de Reconnaissance des Formes et Intelligence Artificielle, Paris, 29 nov-ler déc 1989, pp. 387-400.
- [83] K. Matrouf, J.L. Gauvain, F. Néel, and J. Mariani, "Adapting Probability-Transitions in DP Matching Process for an Oral Task-Oriented Dialogue," IEEE Proc. ICASSP'90, Albuquerque, April 1990.
- [84] K. Matrouf, J.L. Gauvain, F. Néel, and J. Mariani, "An Organ Task-Oriented Dialogue for Air-Traffic Controller Training," SPIE's 1990 Technical Symposium on Optical Engineering and Photonics in Aerospace Sensing, Applications of Artificial Intelligence, Orlando, 16-20 April 1990.
- [85] E. McLarnon, "A Method for Reducing the Frame Rate of a Channel Vocoder by Using Frame Interpolation," Proc. ICASSP'78, pp. 458-461 (Washington, D.C.).
- [86] R. Moore, F. Pereira, and H. Murveit, "Integrating Speech and Natural-Language Processing," Proc. February 1989 DARPA Speech and Natural Language Workshop, pp. 243-247, February 1989.
- [87] R.K. Moore, "The NATO Research Study Group on Speech Processing: RSG10," Proc. Speech Tech 86 (Media Dimensions), pp. 201-203, April 1986.
- [88] J.C. Nagengast, "The STU-III: Narrowband Speech Comes of Age," Proc. Military Speech Tech 1987 (Media Dimensions), November 1987.
- [89] National Research Council CHABA panel on Removal of Noise from a Speech/Noise Signal, "Removal of Noise from Noise-Degraded Speech Signals," Committee on Hearing, Bioacoustics, and Biomechanics (CHABA), National Research Council, published by National Academy Press, 1989.
- [90] J.A. Naylor and S.F. Boll, "Techniques for Suppression of an Interfering Talker in Co-Channel Speech," IEEE Proc. ICASSP'87, pp. 205-208, April 1987.
- [91] F. Néel, K. Matrouf, J.L. Gauvain, and J. Mariani, "Reconnaissance Vocale et Application en Aéronautique," 57ème Congrès de l'ACFAS, Montréal, 17 Mai 1989, pp. 37-42.
- [92] L. Netsch, A. Smith, G. Doddington, and P. Rajasekaran, "Robust Recognition of Speech Under Stress and in Noise," in Proc. National Electronic Convention (NAECON) 1988.
- [93] J.A. Nichols, "Automatic Speech Recognition and Synthesis for Air Traffic Control Trainers," Proc. Military Speech Tech 1988 (Media Dimensions), November 1988.
- [94] D. O'Shaughnessy, *Speech Communication - Human and Machine* (Addison-Wesley) 1987.

- [95] D. O'Shaughnessy, "Speech Enhancement Using Vector Quantization and a Formant Distance Measure," *IEEE Proc. ICASSP'88*, pp. 549-552, April 1988.
- [96] G.A. Osga and E. Craighill, "Man-Machine Interface for C³ Decision Support," *Command, Control, and Communication Technology Assessment: Conf. Proc.*, Defense Communications Agency, November 1986, pp. IV-12-IV-18.
- [97] D.S. Pallett, "Speech Results on Resource Management Task," *Proc. of the February 1989 DARPA Speech and Natural Language Workshop (Morgan-Kaufmann Publishers)*, pp. 18-24, February 1989.
- [98] D.B. Paul and P.E. Blankenship, "Two Distance Measure-based Vocoder Quantization Algorithms for Very-Low-Data Rate Applications: Frame-Fill and Spectral Vector Quantization," in *Proc. IEEE Int. Conf. Commun.*, pp. 1-6, June 1982.
- [99] D.B. Paul, "An 800 bps Adaptive Vector Quantization Vocoder Using a Perceptual Distance Measure," *Proc. ICASSP'83*, pp. 73-76 (Boston).
- [100] D.B. Paul, R.P. Lippmann, Y. Chen, C.J. Weinstein, "Robust HMM-Based Techniques for Recognition of Speech Produced Under Stress and in Noise," *Proc. DARPA Speech Recognition Workshop, February 1986*; also published in *Speech Tech '86 Conf. Proc.*, April 1986.
- [101] D.B. Paul and E.A. Martin, "Speaker Stress-Resistant Continuous Speech Recognition," *Proc. ICASSP'88*, New York City.
- [102] D.B. Paul, "A CSR-NL Interface Specification," *Proc. October 1989 DARPA Speech and Natural Language Workshop (Morgan-Kaufmann)*, pp. 203-214, October 1989.
- [103] B. Plutchak, "Voice Operated Status Boards for the Carrier Air Traffic Control Center," *Proc. Military Speech Tech 1986*, pp. 17-20, 1986.
- [104] A. Poggio, et al., "CCWS: A Computer-based, Multimedia Information System," *IEEE Computer Society Magazine*, Vol. 18, No. 10, pp. 92-105, October 1985.
- [105] D. Pondaco, "Voice Recognition in Army Helicopter Operational Environments," *AVIOS Conf. Proc.*, September 1989.
- [106] G.K. Pooch, "Experiments with Voice Input for Command and Control: Using Voice Input to Operate a Distributed Computer Network," *Naval Postgraduate School Technical Report NPS55-80-016*, April 1980.
- [107] G.K. Pooch and E.F. Roland, "A Feasibility Study for Integrated Voice Recognition Input into the Integrated Information Display (IID)," *Technical Report NPS55-84-008-PR*, Naval Postgraduate School, Monterey, CA, 1984.
- [108] P. Price, W. Fischer, J. Bernstein, and D. Pallett, "The DARPA Resource Management Database for Continuous Speech Recognition," *Proc. ICASSP'88*, New York City, April 1988.
- [109] S.R. Quackenbush, T.P. Barnwell, and M.A. Clements, *Objective Measures of Speech Quality*, Prentice-Hall, Englewood Cliffs, NJ, 1988.
- [110] T.F. Quatieri, J.T. Lynch, M.L. Malpass, R.J. McAulay and C.J. Weinstein, "The VISTA Speech Enhancement System for AM Radio Broadcasting," *MIT Lincoln Laboratory Final Technical Report to the United States Information Agency/Voice of America*, January 1990, DTIC ADA219665.
- [111] T.F. Quatieri and R.J. McAulay, "Noise Reduction Using a Soft-Decision Sine-Wave Vector Quantizer," *IEEE Proc. ICASSP'90*, pp. 821-824, April 1990.
- [112] T.F. Quatieri and R.J. McAulay, "Peak-to-RMS Reduction of Speech Based on a Sinusoidal Model," to be published in *IEEE Transactions on Acoustics, Speech and Signal Processing*.
- [113] L.R. Rabiner and B.H. Juang, "An Introduction to Hidden Markov Models," *IEEE Acoustics, Speech and Signal Processing Magazine*, pp. 4-16, January 1986.
- [114] P.J. Rajasekaran, G.R. Doddington, and J.W. Picone, "Recognition of Speech Under Stress and in Noise," *Proc. ICASSP'86*, Tokyo, April 1986.
- [115] L.W. Reed, "Voice Interactive Systems Technology Avionics (VISTA) Program," *Proc. AGARD Conf. No. 329 on Advanced Avionics and the Military Aircraft Man/Machine Interface*, April 1982.
- [116] J. Rodriguez, J.S. Lim, and E. Singer, "Adaptive Noise Reduction in Aircraft Communication Systems," *Proc. ICASSP'87*, Dallas, pp. 169-172.
- [117] R.C. Rose and D.A. Reynolds, "Text Independent Speaker Identification Using Automatic Acoustic Segmentation," *IEEE Proc. ICASSP'90*, pp. 293-296, April 1990.
- [118] A. Rosenhoover, "Integration of Voice I/O on the AFTI/F-16," *Proc. Speech Tech '84 (Media Dimensions)* 1984.
- [119] A. Rosenhoover, "AFTI/F-16 Voice Interactive Avionics," *Proc. IEEE National Aerospace and Electronics Conf., NAECON*, pp. 613-617, 1986.

- [120] S. Roucos, "A Segment Vocoder at 150 b/s," Proc. ICASSP'83, pp. 61-64 (Boston).
- [121] J. Ruth, A. Godwin, and E. Werkowitz, "Voice-Interactive System Development Program," Proc. AGARD Conf. No. 329 on Advanced Avionics and the Military Aircraft Man/Machine Interface, pp. 26-29, April 1982.
- [122] J.C. Ruth, "Application of Voice Interactive Systems," Speech Tech '85, 1985.
- [123] H. Schecter and J. Tierney, "Operational Acceptability of 2.4 Kbps Speech for Tactical Communications," Proc. MILCOM'85, pp. 171-180, Boston, October 1985.
- [124] R. Schwartz and Y.L. Chow, "The N-Best Algorithm: an Efficient and Exact Procedure for Finding the N Most Likely Sentence Hypotheses," IEEE Proc. ICASSP'90, pp. 81-84, April 1990.
- [125] C. Searle, "Voice Aspects of UK JTIDS Integration Flight Trials, Summer 1988," 1989 Speech Test and Evaluation Workshop, RADC/EEV, Hanscom AFB, March 1989.
- [126] C.A. Simpson, C.R. Coler, and E.M. Huff, "Human Factors of Voice I/O for Aircraft Cockpit Controls and Displays," Proc. NBS Workshop on Standardization for Speech I/O Technology, pp. 159-166, March 1982.
- [127] G. Slemon and D. Eames, "Speech Technology Applied to Operational Air Traffic Controller Training Systems," Proc. Military Speech Tech 1986 (Media Dimensions) October 1986.
- [128] K. Smith and W.J.C. Bowen, "The development of Speech Recognition for Military Fast-Jet Aircraft," Proc. Military Speech Tech 1986 (Arlington, VA).
- [129] D.M. Snider, "SCAMP - Single Channel Advanced Man Portable Terminal," Proc. U.S. Army Communications Command Symposium on Space: Supporting the Soldier, Ft. Monmouth, NJ, June 1989.
- [130] B.J. Stanton, L.H. Jamieson, and G.D. Allen, "Acoustic-Phonetic Analysis of Loud and Lombard Speech in Simulated Cockpit Conditions," in IEEE Proc. Int. Conf. Acoustics, Speech and Signal Processing 1988, April 1988.
- [131] B.J. Stanton, L.H. Jamieson, and G.D. Allen, "Robust Recognition of Loud and Lombard Speech in the Fighter Cockpit Environment," in IEEE Proc. Int. Conf. Acoustics, Speech and Signal Processing 1989, pp. 675-678, May 1989.
- [132] H.J.M. Steeneken, "Quality Evaluation of Speech Processing Systems," Proc. AGARD Lecture Series No. 170 on "Speech Analysis and Synthesis and Man-Machine Speech Communication for Air Operations," May 1990.
- [133] R.J. Stubbs and Q. Summerfield, "Separation of Simultaneous Voices," in Proc. of European Conf. on Speech Technology, Edinburgh, Scotland, 1987.
- [134] R.I. Sudhoff, "The ANDVT Program of Equipments," Proc. Military Speech Tech 1987 (Media Dimensions), November 1987.
- [135] R.F. Swider, "Operational Evaluation of Voice Command/Response in an Army Helicopter," Proc. Military Speech Tech 1987 (Media Dimensions), Arlington, VA, November 1987.
- [136] P. Tarnaud, "Integration du Dialogue dans un Environnement Reel Complexe," GALF/GRECO Workshop on Man-Machine Dialogue by Voice, Nancy, October 1984.
- [137] P. Tarnaud, "Applications of Voice I/O in Future Cockpit Design," Proc. Speech Tech '86, pp. 63-65 (New York City).
- [138] P. Tarnaud, "Speech Processing: the Crouzet Research and Development Programme," Proc. IEEE National Aerospace and Electronics Conf., NAECON, pp. 803-812, 1986.
- [139] M.R. Taylor, "DVI and Its Role in Future Avionic Systems," Proc. IFS Conf. on Speech Technology, pp. 113-120, Brighton, UK, October 1984.
- [140] M.R. Taylor, "Voice Input Applications in Aerospace," in *Electronic Speech Recognition*, G. Bristow (Ed.), Collins, pp. 322-248, 1986.
- [141] M.M. Taylor, F. Neel, and D.G. Bouwhis, eds., *The Structure of Multi-Modal Dialogue* (North Holland, Elsevier Science Publishers, 1989).
- [142] J. Tierney and H. Schecter, "Experiences with Narrow-band Speech in Tactical Test Environments," Proc. Military Speech Tech 1987 (Media Dimensions), November 1987.
- [143] R. Turn, A. Hoffman, and T. Lippiatt, "Military Applications of Speech Understanding Systems," RAND Corp. Technical Report R-1434-ARPA, June 1974, (AD-787394/6GA).
- [144] A.P. Varga and K.M. Ponting, "Control Experiments on Noise Compensation in Hidden Markov Model Based Continuous Speech Recognition," ESCA Proc. Eurospeech'89, Paris, September 1989.
- [145] A.P. Varga and R.K. Moore, "Hidden Markov Model Decomposition of Speech and Noise," Proc. ICASSP'90, Albuquerque, pp. 845-848, April 1990.
- [146] G. Vensko, "A Single Board High Performance Recognizer," Proc. Military Speech Tech 1986 (Media Dimensions) October 1986.

- [147] V.R. Viswanathan, C.M. Henry, and A. Derr, "Noise Immune Speech Transduction using Multiple Sensors," Proc. ICASSP'85, Tampa, pp. 712-715.
- [148] C.J. Weinstein and J.W. Forgie, "Experience with Speech Communication in Packet Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 1, No. 6, pp. 963-980, December 1983.
- [149] C.J. Weinstein, D.B. Paul, and R.P. Lippmann, "Robust Speech Recognition Using Hidden Markov Models: Overview of a Research Program," MIT Lincoln Laboratory, Lexington, Mass., Technical Rep. 875, (February 1990).
- [150] R.M. Weischedel, R.J. Bobrow, D.A. Ayuso, and L. Ramshaw, "Portability in the JANUS Natural Language Interface," Proc. DARPA Speech and Natural Language Workshop (Morgan-Kaufmann), February 1989, pp. 112-117.
- [151] R. Weischedel, J. Carbonell, B. Grosz, W. Lehnert, M. Marcus, R. Perrault, and R. Wilensky, "White Paper on Natural Language Processing," Proc. October 1989 DARPA Speech and Natural Language Workshop, pp. 481-493, October 1989.
- [152] M. Weiss and E. Aschkenasy, "Study and Development of the INTEL Technique for Improving Speech Intelligibility," Technical Report RADC-TR-75-108, Rome Air Development Center, Griffiss AFB, NY, 1975.
- [153] M. Weiss and E. Aschkenasy, "The Speech Enhancement Advanced Development Model," Technical Report RADC-TR-78-232, Rome Air Development Center, Griffiss AFB, NY, 1978.
- [154] E. Werkowitz, "Speech Recognition in the Tactical Environment: The AFTI/F-16 Voice Command Flight Test," Proc. Speech Tech '84 (Media Dimensions) 1984.
- [155] C.J. Westerhoff and L. Reed, "Voice Systems in U.S. Army Aircraft," *Speech Technology Magazine*, February/March 1985.
- [156] R.G. White, "Automatic Speech Recognition as a Cockpit Interface," AGARDograph No. 272, Advances in Sensors and their Integration into Aircraft Guidance and Control Systems, 1983.
- [157] R.G. White, "Speaking to Military Cockpits," RAE Tech. Memo FS(B) 671, June 1987.
- [158] D.T. Williamson, "Flight Test Results of the AFTI/F-16 Voice Interactive Avionics Program," Proc. American Voice I/O Society (AVIOS) 87 Voice I/O Systems Applications Conf., Alexandria, VA, pp. 335-345, October 1987.
- [159] J.P. Woodard and E.J. Cupples, "Selected Military Applications of Automatic Speech Recognition," *IEEE Communication Magazine*, Vol. 21, No. 9, pp. 35-44, December 1983.
- [160] S.J. Young, "Competitive Training in Hidden Markov Models," Proc. ICASSP'90, Albuquerque, pp. 681-684.
- [161] M.A. Zissman, "Co-Channel Talker Interference Suppression," Ph.D. Thesis, MIT Department of Electrical Engineering and Computer Science, January 1990.
- [162] M.A. Zissman, C.J. Weinstein, and L.D. Braida, "Automatic Talker Activity Labeling for Co-Channel Talker Interference Suppression," *IEEE Proc. ICASSP'90*, pp. 813-816, April 1990.
- [163] V. Zue, J. Glass, D. Goodine, H. Leung, M. Phillips, J. Polifroni, and S. Seneff, "The VOYAGER Speech Understanding System: Preliminary Development and Evaluation," *IEEE Proc. ICASSP'90*, pp. 73-76, April 1990.