AD-A197 219

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(1)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFIT/CI/NR 88- 171	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) VOICE RECOGNITION AND ARTIFICIAL INTELLIGENCE IN AN AIR TRAFFIC CONTROL ENVIRONMENT		5. TYPE OF REPORT & PERIOD COVERED MS THESIS
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) ROBERT FRANCIS HALL		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS AFIT STUDENT AT: ARIZONA STATE UNIVERSITY		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE 1988
		13. NUMBER OF PAGES 107
14. MONITORING AGENCY NAME & ADDRESS (If different from Controlling Office) AFIT/NR Wright-Patterson AFB OH 45433-6583		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

DISTRIBUTED UNLIMITED: APPROVED FOR PUBLIC RELEASE

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

SAME AS REPORT

18. SUPPLEMENTARY NOTES

Approved for Public Release: IAW AFR 190-1
LYNN E. WOLAVER Lynn Wolaver 8 Aug 88
Dean for Research and Professional Development
Air Force Institute of Technology
Wright-Patterson AFB OH 45433-6583

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

ATTACHED

DTIC
ELECTE
S AUG 1 7 1988 D
H

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

88 8 10 08

Abstract

As the growth of air carriers continues, this nation's Air Traffic Control system is further strained. To date, most efforts at improving this system have been in the form of adding more men and machinery to an ever expanding control network. It is the purpose of this study to determine the areas within the ATC system which may benefit from the introduction of voice recognition and artificial intelligence, and to develop prototype equipment from which experiments can be carried out.

This study expands on work accomplished by Russell W. White and Donald L. Parks in their final technical report submitted in response to NASA Contract NAS1-17367, "Study To Determine Potential Flight Applications and Human Factors Design Guidelines for Voice Recognition and Synthesis Systems" by considering many of these guidelines in an Air Traffic Control environment. A procedure is presented to aid in the development of grammar definition programs for voice recognizers through the use of a sentence diagramming technique. Test results for the performance of prototype recognition equipment are given. Applications for data transmission and storage, expert systems, and switching devices are discussed.

VOICE RECOGNITION AND ARTIFICIAL INTELLIGENCE
IN AN
AIR TRAFFIC CONTROL ENVIRONMENT

by

Robert Francis Hall

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science in Engineering

ARIZONA STATE UNIVERSITY

May 1988

VOICE RECOGNITION AND ARTIFICIAL INTELLIGENCE

IN AN

AIR TRAFFIC CONTROL ENVIRONMENT

by

Robert Francis Hall

has been approved

May 1988

APPROVED:

______________________ Chairperson

____

____

Supervisory Committee

ACCEPTED:

____

Department Chairperson

____

Dean, Graduate College

Abstract

As the growth of air carriers continues, this nation's Air Traffic Control system is further strained. To date, most efforts at improving this system have been in the form of adding more men and machinery to an ever expanding control network. It is the purpose of this study to determine the areas within the ATC system which may benefit from the introduction of voice recognition and artificial intelligence, and to develop prototype equipment from which experiments can be carried out.

This study expands on work accomplished by Russell W. White and Donald L. Parks in their final technical report submitted in response to NASA Contract NAS1-17367, "Study To Determine Potential Flight Applications and Human Factors Design Guidelines for Voice Recognition and Synthesis Systems" by considering many of these guidelines in an Air Traffic Control environment. A procedure is presented to aid in the development of grammar definition programs for voice recognizers through the use of a sentence diagramming technique. Test results for the performance of prototype recognition equipment are given. Applications for data transmission and storage, expert systems, and switching devices are discussed.

Acknowledgments

I would like to thank the men and women of the 1922 Communications Squadron at Williams AFB for their help and patience throughout this project. A special thanks goes to Captain Doug Volkman, who was always open to new ideas, and A1C Ed Ferguson, who was reciting aircraft callsigns in his sleep by the end of this project. Thanks also go to Dr. Hewitt Young, whose two years of instruction and advice led to this project.

Table Of Contents

List of Tables

List of Figures

Chapter One

INTRODUCTION

The rapid growth of air carrier, general aviation, and military traffic has strained this nation's Air Traffic Control (ATC) system. The symptoms of this strain appear as controller fatigue, low controller moral, and the occasional creation of a hazardous situation caused by human error.

The current method employed to improve the ATC system has been in the form of increasing its air traffic handling capacity by adding more machinery and manpower. Thus, machines with greater processing power and more humans are coupled into a man-machine system which is destined to continually grow. Little has been done in finding new forms of technology to increase the joint efficiency of man and machine.

Two relatively new technologies which could create a path towards greater system efficiency are the technologies of voice recognition and artificial intelligence. With greater system efficiency, less controller fatigue and better air safety are expected. Where to apply these technologies, in what form, and how deep these technologies can be integrated into the ATC system are questions which deserve inquiry. This research details a method to answer these questions, develops prototype equipment from which to experiment, and establishes a basis from which other research efforts may be launched.

A review of literature indicates that current efforts at applying voice recognition in flight operations are centered around pilot task improvement and special projects such as the space shuttle. While these efforts are certainly important, the controller environment can be just as demanding, stressful, and important. It is within this light that this study is launched.

The Research Concept

It is the purpose of this study to determine the areas within the ATC system which can benefit from the introduction of voice recognition and artificial intelligence, and to develop prototype equipment from which experiments can be carried out. In addition, the research effort seeks to answer the following questions:

- What is the structure and characteristics of the ATC language and will this language lend itself towards machine processing?
- Is voice data collection a feasible method to collect controller information under current technology?
- Once collected, can controller information be processed in a useful and beneficial way?
- Can artificial intelligence be used in conjunction with information collection to provide an advanced machine interrogation system with applications such as the warning of the existence of hazardous air traffic situations?

Voice recognition can be applied to many facets of the ATC system. To help with this investigation, this study considered certain guidelines proposed and results obtained by Russell W. White and Donald L. Parks in their final technical report submitted in response to NASA Contract NAS1-17367, "Study To Determine Potential Flight Applications and Human Factors Design Guidelines for Voice Recognition and Synthesis Systems."

The White and Parks study surveyed voice recognition and synthesis equipment up to 1984 and determined its applicability to flight and simulator operations. Their guidelines for implementing voice recognition and synthesis systems in aircraft and simulator cockpits appears in Appendix A. While the guidelines were designed for applications in the air, they can serve as a starting point for an analysis of ATC voice recognition applications.

In their study, White and Parks note that highest benefit in voice recognition systems seem to come from an enhanced interrogation capability between man and a complex machine system as opposed to the use of voice recognition as a switching or selection device. This study investigates the use of voice recognition as both an interrogation system and as a selection/switching system. For use as an interrogator, postprocessing of the recognized vocabulary is necessary.

Air Traffic Control Language

The ATC language lends itself towards machine processing. The language is very precise, with a particular word order, and is spoken by users who receive training in its exact execution. The establishment of an ATC voice recognition vocabulary is of primary concern.

An analysis of the ATC language was conducted. This analysis consisted of the study of controller voice tapes and experiments conducted with prototype voice recognition equipment. An ATC voice recognition language was developed within contextual frames to define the number and depth of the language's context. With this information, a grammar definition file (developed using a proprietary symbolic notation system from Interstate Voice Products) was created for processing by prototype voice recognition equipment.

Voice Recognition

Initial considerations in this area were on the effects of the ATC environment on the application of voice recognition. Environmental noise, message syntax, and message vocabulary have a major effect on recognition accuracy. Environmental noise impacts the recognition system by possibly interfering with the voice recognizer. Message syntax and vocabulary impact the recognition system by ultimately determining the grammar definition language and recognizer processing time.

The pragmatic application of a voice recognition system using off-the-shelf equipment was of major consideration throughout the study. This application was based on the following assumptions:

- for current voice recognition systems to achieve high benefit under current technology, the system must be based on speaker dependent voice patterns.
- the system must be capable of at least connected-word recognition.
- each controller station will possess voice recognition capability.

The preceding assumptions point toward a low-cost PC-based recognition system capable of being networked with a host computer. The individual controller's PC voice recognition unit would interpret the controller's speech and send it to a host computer, which would process the data. The following PC-based recognition boards or units appeared promising and their manufacturers were contacted to arrange for on-loan or rental prototype equipment:

- TI-Speech System, Texas Instruments.
- VocaLink CSRB, Interstate Voice Products.
- Voice Developer System, Verbex Voice Industries Corporation.
- Coretechs VET 3 Speech Terminal, Scott Instruments Corporation.

The prototype equipment was tested at the Air Traffic Control facility at Williams Air Force Base. The recognition unit used in these tests was first brought to my attention as the Voice Developer System distributed by Verbex Voice Industries Corporation. The unit is now distributed by Interstate Voice Products as the VocaLink® S4000. Interstate Voice Products also distributes the VocaLink® CSRB and SRB-LC recognition systems as shown above.

The VocaLink® S4000 is a state-of-the-art standalone voice recognition unit capable of connected-speech recognition. The unit provides grammar control through the use of a symbolic notation system. The grammar defines the phrase pattern-matching abilities of the unit. User specified translation tables allow the unit to communicate with a host computer running voice postprocessing programs. The October 27, 1987 issue of *PC Magazine* reviews the unit under the name of Verbex Voice Developer System.

Testing consisted of determining the recognition accuracy of a particular grammar in an operational Air Traffic Control environment. Recognition grammars were first developed and tested in an office setting until the grammar was refined. Once refined, the grammar was tested in an operational setting using a qualified Air Traffic Controller as the subject.

Artificial Intelligence

The application of artificial intelligence in an ATC voice recognition system is best suited for the postprocessing of recognition information. The VocaLink® S4000 recognition unit allowed its output to be defined through the use of a translation table. This feature allows for the configuration of input to a postprocessing program. Of particular concern is the use of artificial intelligence programs to store, display, and manipulate voice data in a user-friendly way. This study used Borland's Turbo PROLOG language to develop standalone programs to display, store, and manipulate voice recognition data.

In addition, the ATC language was defined using a method to facilitate artificial intelligence processing. Contextual sentence diagrams were developed to define the context in which each word in a particular vocabulary is uttered. For instance, a typical Air Traffic Controller could define his vocabulary universe through less than five verbs. These five or fewer verbs could have several noun phrases associated with each verb. The controller's context of speech can be determined by his or her particular verb usage. This process parallels the many verb-centered semantic theories found in linguistics.

THE NECESSITY OF VOICE

A written language provides a portable, exact and manipulable medium for conveying thought. The words of a written language may be reproduced and transmitted, they usually have an exact meaning, and they can be sorted, counted, or otherwise processed. Common practice dictates that words should "stand on their own" having little requirement for background knowledge.

However, a written language requires effort. Tools must be devised to capture letters on paper. The author must slave away ensuring that spelling, punctuation, and grammar are all properly accomplished. Complicated machinery must be used in order to process the written word.

A verbal language provides a convenient and rich way to express thought. It requires no special apparatus and its rich tonal properties allow all types of messages to be expressed, even when the same words are used. But, verbal language is difficult to process. It can't be sorted or manipulated without first being converted to a written language except through great mental effort. When converted to written language, verbal language loses its tonal qualities and its true meaning may not be conveyed. Thus, written language provides precision while verbal language provide convenience and variety.

Until recently, it was generally believed that the processing of information would always have to begin from the input of a written language. Whether english words for a word processing program or cryptic words for a programming language, the input was always in a symbolic written form. This was primarily due to the lack of a verbal input interface for information processing machinery (no verbal keyboard) and the requirement for input precision.

Today, the verbal input interface is a reality, but we are still plagued by a lack of precision in verbal input. This does not mean that a user cannot be trained to speak precisely within predefined constraints, but that humans are used to speaking with very few constraints. However, in some applications such as the Air Traffic Control environment, verbal precision is already imposed. Within these applications an investigation into the use of a verbal interface to capture and process information, and the benefits which may be provided by such an interface deserve inquiry.

Modalities

The human is constantly attempting to learn more about his or her environment either for survival or curiosity. As such, we continually receive information from our environment. This information is perceived by the human senses of vision, hearing, touch, taste and smell. Each sense provides a mode through which information may travel.

The visual mode of information encompasses perceptions from the flash of a light, to the printed word you are now reading. The auditory mode of information encompasses perceptions from the roar of a lion, to the "ticking" sound your auto mechanic hears when he recommends that you replace the valves in your car engine. Touch information encompasses perceptions from a blow to the face, to the bump a skilled craftsmen may feel when he runs his hand over a newly planed piece of wood. The modality of taste can encompass the sour taste a survivalist may use to determine if a plant is edible, to the rare taste a French chef may be checking for in a soufflé. Finally, the perceptions of smell may range from the odor of "methyl mercaptan" commonly mixed with natural gas, to the fragrance of a friend's perfume or cologne. These sense modalities provide us with the information we need in order to interact with our environment.

Unfortunately, through experience we have found that the use of these modalities can be quite limited. Some people cannot perceive certain stimuli that others can detect and at other times the stimulus may be short in duration. In these cases, we have found it possible to relate the experience to another person by expressing the stimulus in another modality. Thus, the fisherman can't show you the "one that got away", but he can tell you how big it was. The fisherman mixes modalities of information in order to relate his experience.

The mixing of modalities in order to relate experience can cause problems in the form of misinformation:

- Information may be added.
- Information may be lost.
- The full expression of the stimuli may not be possible in another modality.

In the case of the fisherman, what may start as a twelve inch perch, becomes a whale.

For historical and transmission purposes, information is constantly being changed from one modality to another. This can be seen in the cases of a secretary typing a dictated letter, a physician writing the symtoms expressed by a patient in a report, or an Air Traffic Controller typing in information received from a pilot's radio message. The information is constantly being "processed". Usually this processing is to translate the information into a written language so that it may be retrieved at a later time or different location. Thus all the information received through five senses is usually channeled to the one **visual** modality through the use of written or cryptic languages.

The use of **verbal** information input can eliminate some of the difficulties associated with information presented through mixed modalities. Information obtained directly from the speaker within a precisely defined grammar can provide the precision needed for information input without the possibility of misinformation associated with mixed modality information processing and without the additional effort needed to convert the information into written form.

The determination of what a user is saying as he or she says it, provides for immediate information flow. Once obtained this information can be displayed, stored, or manipulated (compared). In the case of an expert system application, it is possible to compare speech output against a predefined rule base and to provide immediate feedback to the user if an unsafe condition is determined.

Outline of Thesis

Chapter One provided you with an introduction to the research effort and an explanation as to why voice recognition may be beneficial. Chapter Two continues to introduce you to the theory and use of voice recognition by examining the human and machine voice recognition systems.

Chapter Three describes the analysis of the Air Traffic Control language, the contextual sentence diagrams, and the grammar definition programs used to define the Air Traffic Control language. Chapter Three also introduces a method to aid in the establishment of a verbal grammar.

Chapter Four describes the method used to test the feasibility of integrating a voice recognition system into the Air Traffic Control environment. The implementation of the ground controller grammar on the VocaLink® S4000 voice recognition unit is discussed.

Chapter Five explores possible applications of voice recognition in an Air Traffic Control environment. The design and suitability of the applications of switching, data collection and transmission, and expert systems are discussed. Chapter Six provides a conclusion of the study and suggestions for further research in this area.

Chapter Two

This chapter provides the reader with a review of human and machine voice recognition systems. It begins with an explanation of the human system covering its three elements: the speaker, the language, and the listener. This is followed by an examination of current voice recognition technology.

THE HUMAN VOICE RECOGNITION SYSTEM

The human recognition system is comprised of three elements: the speaker, the language and the listener. These three elements form a complicated link in voiced communication. This link is depicted in figure one. The failure of any one element will result in a failure to communicate. To understand the human voice recognition system, a thorough understanding of each element is necessary.

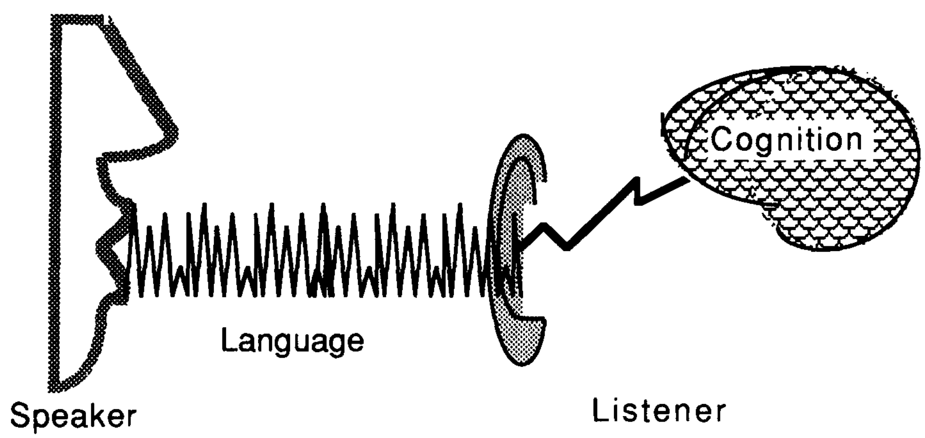

Figure 1. Human voice recognition system.

The Speaker

The expression of a thought through voice, requires a series of physiological events with the ultimate result being an utterance. These events include the compression of the lungs, the vibration of the vocal cords, and manipulation of the vocal tract.

The primary function of the lungs is to supply our body with oxygen and to get rid of carbon dioxide. Each lung is a mass of spongy tissue surrounded by an air tight membrane called the pleura. As we breathe, a dome shaped muscle called the diaphragm, located below the lungs, contracts. This contraction creates a vacuum around the lungs and the lungs inflate. When the diaphragm relaxes, the lungs contract and air is forced out of the lungs as shown in figure two.

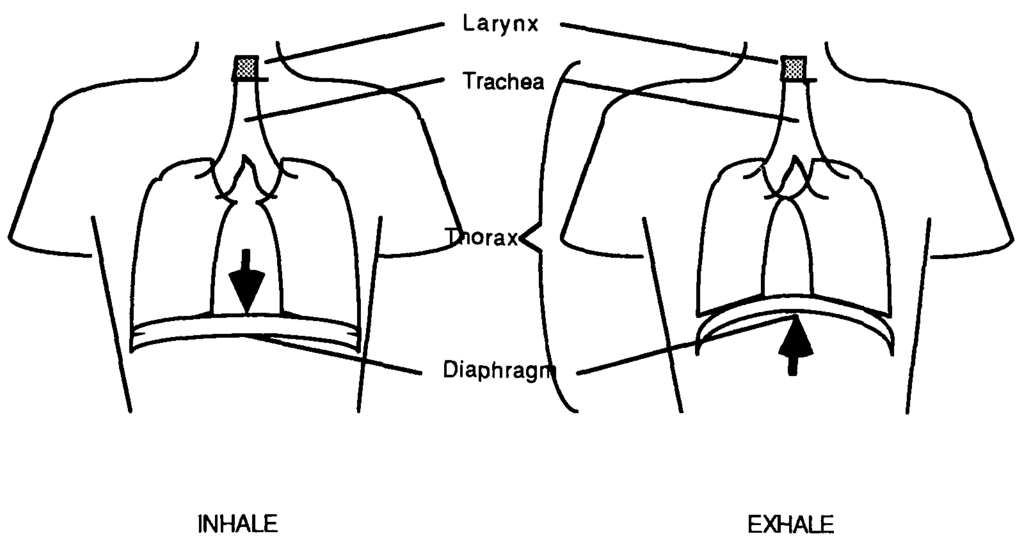

Figure 2. Human lungs.

At times, it may be necessary to create additional force in exhaling in order to cough or blow. In this case, the abdominal muscles can be used to apply pressure to the bottom of the diaphragm. The end result is that the lungs provide us with the means to pressurize the thorax. It is this pressurized air which is used to vibrate stretched flaps of fleshy tissue located in the larynx and called the vocal cords.

The larynx is located at the top of the trachea. It contains a mass of cartilages and muscles used to control the vocal cords. Each cord is attached to thyroid cartilage in the front and arytenoid cartilage in the back. A set of muscles attaches the arytenoid cartilages to the cricoid cartilage as shown in figure three. These muscles move the vocal cords together and apart.

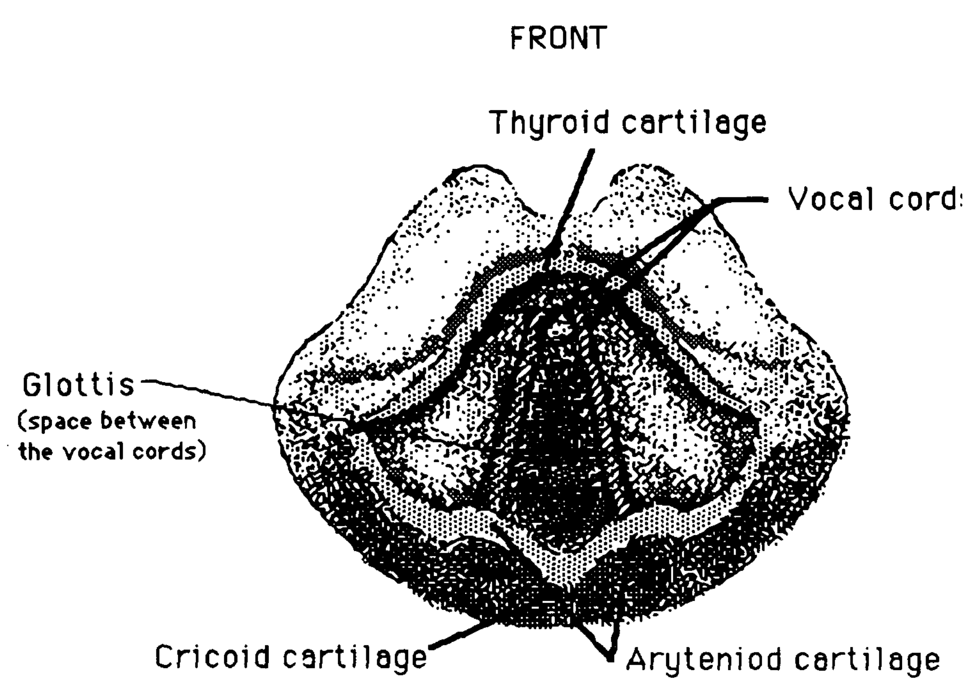

Figure 3. Human vocal cords (Kunz p. 351).

The vocal cords are used to seal off the trachea during swallowing and coughing, and to vibrate during the production of sound. The frequency of the vibration is dependent on the mass and tension of the cords ranging from 80 to 400 vibrations per second. (Kantowitz p. 274) The vibration of the vocal cords can be accomplished in different ways including:

- **phonation**: the vocal cords are drawn fully together and the cords emit a continuous vibration.
- **whispering**: the vocal cords are drawn nearly together and the cords emit a whiteband noise along with a continuous vibration. (Parsons p. 65)

The vocal tract consists of the air passageway past the larynx. This area consists of the laryngeal pharynx, oral pharynx, nasal cavity, and oral cavity. The vibrations emitting from the vocal cords are resonated and articulated using a valve-like system of appendages. For instance, the uvula is capable of sealing off the nasal cavity and the tongue is capable of obstructing airflow over a wide range of the oral cavity (see figure four).

The obstruction of airflow through the vocal tract can have three effects on voice:

- **frication**: the vocal tract is constricted at some point resulting in a broadband noise characterized by the point of constriction.
- **compression**: the vocal tract is completely obstructed allowing pressure to build up which when released results in a short noise burst.
- **vibration**: air is forced through a closure other then the vocal cords such as the tongue or lips (Parsons p.66).

The vocal tract can be modeled as a 17cm oral tube with a second 13cm nasal tube branching off near its middle. The velum acts as a valve to permit air to pass through the nasal tube. Once excited by the vocal cords, these tube structures develop resonant

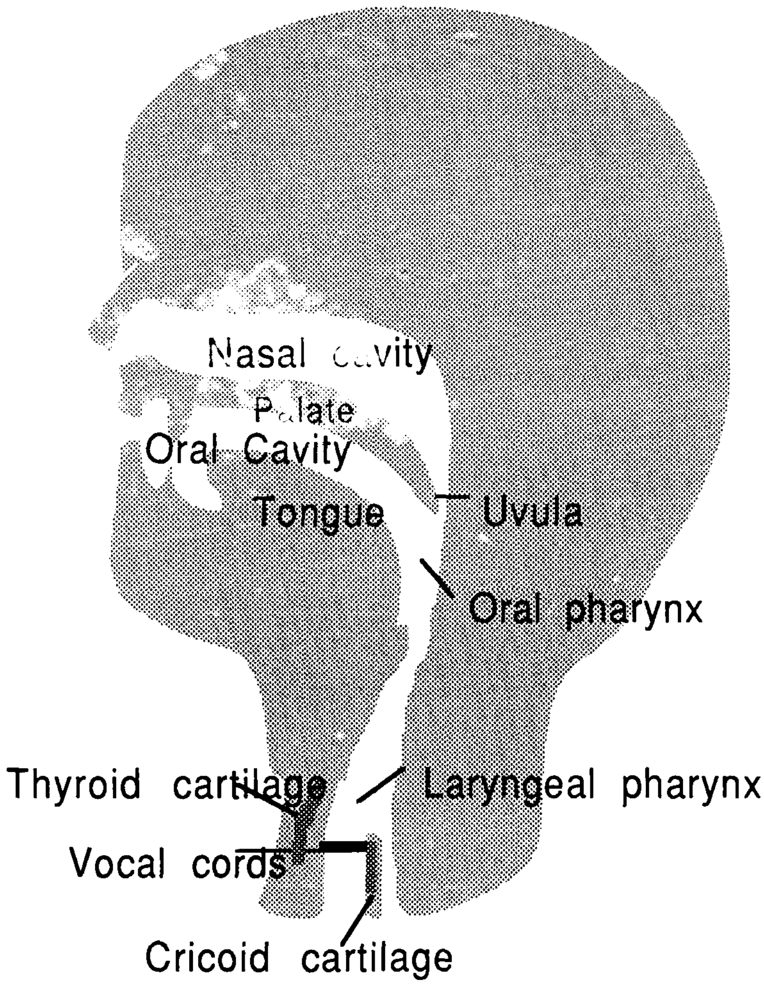

Figure 4. Vocal tract.

frequencies called formants. Typical format frequencies for an adult male are: F1=200-800Hz, F2=600-2800Hz, and F3=1300-3400Hz (Parsons, p. 104). The result is a rich assortment of sounds with a frequency range of 80 to 4000hz for vowel sounds and up to 8000hz for consonants.

The Language

The rich assortment of sounds in voice can be classified in a number of ways including the allophone, phoneme, diphone, syllable and word. The allophone is the finest category of voice. It distinguishes differing sets of phonemes having the same features. An example of its use would be found in distinguishing between the two phrases "he lit" and "heal it". Both phrases would be symbolized in the International Phonetic Alphabet by "hilɪt" . The allophone form of these phrases is "hilɪt^h" and "hi**l**ɪt^h", respectively. The difference is found in the light [l] in the first case and a dark [**l**] in the second. A syllable or word boundary will occur before the light [l] and after the dark [**l**] (Shoup, p. 128). Allophones number in the thousands.

The phoneme has been defined as "the smallest speech sound that can change the meaning of a word" (Kantowitz, p. 274) and "the total collection of allophones that function similarly and do not make meaningful distinctions among themselves within a given language" (Shoup, p. 128). It is basically a set of language sounds which are heard the same within a particular language. The phoneme is commonly used in dictionaries to show the pronunciation of words. Phonemes are a manageable set of sounds, in English usually numbering around 47. Table one depicts the Phonetic Arpabet designed for computer output (Parsons,p. 112).

A diphone is the sound of a pair of phonemes found in the transition from a verb to a consonant or a consonant to a verb. This category can be useful in voice recognition because much of the acoustical information which identifies a consonant is found in this transition sound. Diphones number in the thousands.

TABLE 1

Arpabet Computer Phonetic Alphabet

Arpabet	Example	Arpabet	Example	Arpabet	Example
i	heed	X	roses	m	mom
I	hid	p	pop	n	noon
e	hayed	b	bob	G	ringing
E	head	t	tug	l	lulu
@	had	d	dug	L	battle
a	hod	k	kick	M	bottom
c	hawed	g	gig	N	button
o	hoed	f	fife	F	batter
U	hood	v	verve	Q	glottal stop
u	who'd	T	thick	w	wow
R	heard	D	those	y	yoyo
x	ahead	s	cease	r	roar
A	bud	z	pizzaz	C	church
Y	hide	S	mesh	J	judge
W	how'd	Z	measure	H	where
O	boy	h	heat		

A syllable is an uninterrupted utterance of a word or part of a word. Due to the uninterrupted nature of syllables, they tend to be easily identified, although sometimes the boundaries between syllables are difficult to determine. For example, the syllables of "common" could be split as "kam" +"ən" or "ka" +"mən". Although fewer in number then allophones or diphones, there are hundreds of syllables.

Words form the category of voiced sounds which carry meaning. With correct pronunciation, single words are usually unambiguous. However, when words are spoken rapidly together, the words can become "slurred". This occurs when the end of one word and beginning of the next word becomes difficult to determine.

The Listener

The listener is characterized by the physiological event of hearing and the mental act of cognition. Shared meaning can only occur if the message is both heard and understood. The mental act of cognition occurs within a certain context. This context can be thought of as the mental set of the listener.

The human ear consists of three parts: the outer, middle and inner ear as shown in figure five. The outer ear is made up of the *pinna*, the auditory canal, and the tympanic membrane. The pinna provides protection and directional cues for the ear. The auditory

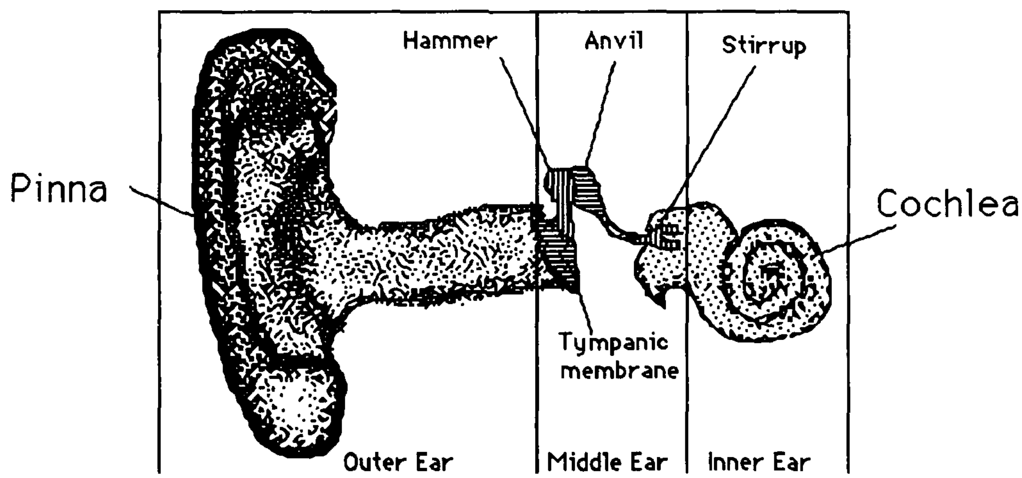

Figure 5. Human ear.

canal is a tube about 2.7 cm long with a resonant frequency around 3000Hz. The tympanic membrane or eardrum separates the outer and middle ear. As an acoustical couple, it vibrates in response to a sound and mechanically transmits the vibration to the middle ear.

The middle ear contains three bones: the hammer, anvil, and stirrup. These three bones provide an acoustical link from the outer ear to the inner ear. The hammer attaches to the tympanic membrane, the stirrup attaches to the inner ear, and the anvil connects these two bones. This system of bones provides an efficient means of transmitting acoustical energy to the fluid-filled inner ear and sound amplitude limiting (Parsons, p. 68).

The inner ear contains the cochlea. The cochlea is a spiraled tube which is partitioned by two membranes called the basilar and tectorial membranes. These membranes along with the organ of Corti form a mechanism which translates the vibration of the inner ear fluid into nerve impulses. The determination of the frequency of a sound appears to be a combination of the location of membrane stimulation and the frequency of stimulation.

Determining if the sound heard is part of a vocal language appears to be dependent on the preconceived phonological rules of the listener. If the initial sound is a recognizable phoneme, it is used in limiting the number of possible words the sound could be a part of. As more sounds become available, the listener determines a word. The determination of a word, signals the start of a new word and the process reiterates. As words become recognized, information is extracted and this information is used to guide the analysis of later words (Parsons, p. 77).

The extraction of information occurs within a context. Context is present because the human continually analyzes the information he or she receives. This analysis takes the form of a working model which guides the human's search for and expectation of information. Figure six portrays this analysis mechanism.

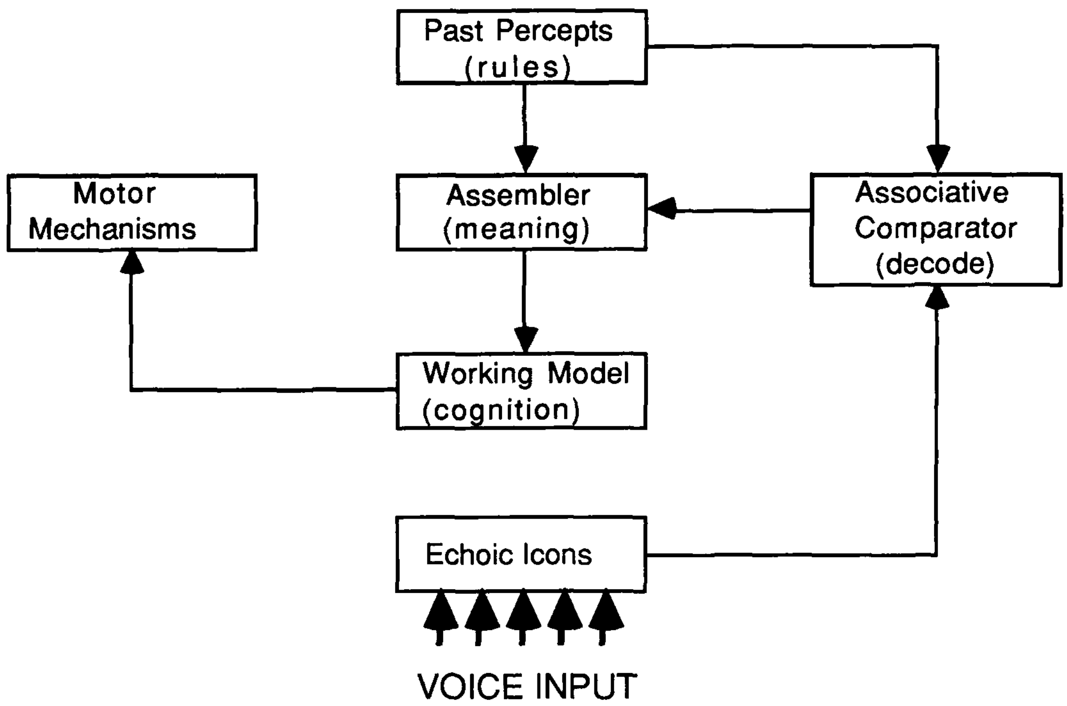

Figure 6. Voice analysis mechanism (Sowa, p. 35).

The neural excitation produced by voice has a persistence which exhibits itself as an echoic icon to the brain. The brain has a predisposed ability to decode voice echoic icons. This decoding occurs at the speech-processing center of the brain located in the left hemisphere (Parsons, p. 77). Here, the icons are compared with phonological rules developed from past perceptions. As words become decoded, they are assembled in order to extract meaning. The meaning of a group of words is developed into a working model which we continually update as new information arrives. The working model guides our search for additional information through our motor mechanisms.

MACHINE VOICE RECOGNITION SYSTEMS

With a background on how the human recognition system works, it is possible to look at machine voice recognition technology from the vantage of knowing how the ultimate voice recognizer

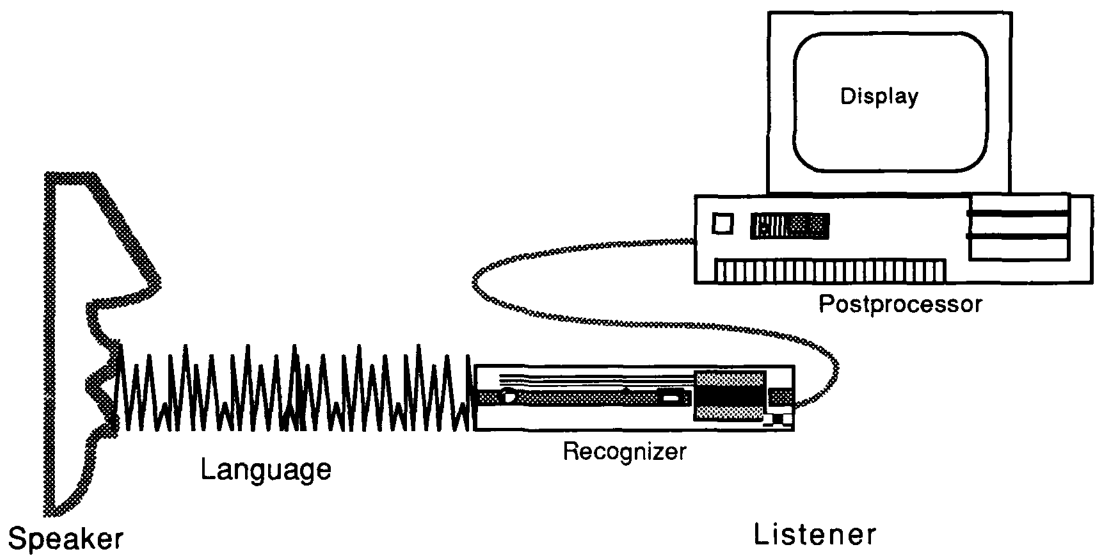

Figure 7. Machine voice recognition system.

should work. As seen in figure seven, the speaker and language elements of the machine voice recognition system remain unchanged, only the listener is different. The new listener element must be capable of perceiving and understanding the voice language.

The Recognizer

The recognizer is presented with a language acoustical signal. This is usually accomplished by having the speaker talk into some type of microphone. Many types of microphone arrangements are possible, the arrangement used with the VocaLink® S4000 recognition unit was a Shure VR230A head-worn dynamic microphone. The head-worn microphone allows for good voice pickup with reasonable speaker mobility. Upon receiving an acoustical signal, the recognition unit must determine what to do with it.

The recognition unit may do three things with the signal. It may correctly identify the signal as a word in its defined grammar. It may mistakenly identify the signal as another word in its defined grammar. It may not respond at all. In order to correctly identify the signal the recognition unit must obtain language information from the acoustical waveform.

The acoustical waveform carries with it frequency and timing information. This information may be pictorially represented by a spectrogram. A typical spectrogram for the callsign "Awol two two" appears in figure eight. Notice that all sounds are characterized by the excitation of the acoustical waveform at certain and many times multiple frequencies. If one could characterize each vocal sound with a unique frequency or set of frequencies, it would be possible to uniquely define all vocal sounds. Unfortunately this categorization of sound is not very easy due to differences among speakers and even differences for the same speaker at different times, but these frequency categories can provide cues for the

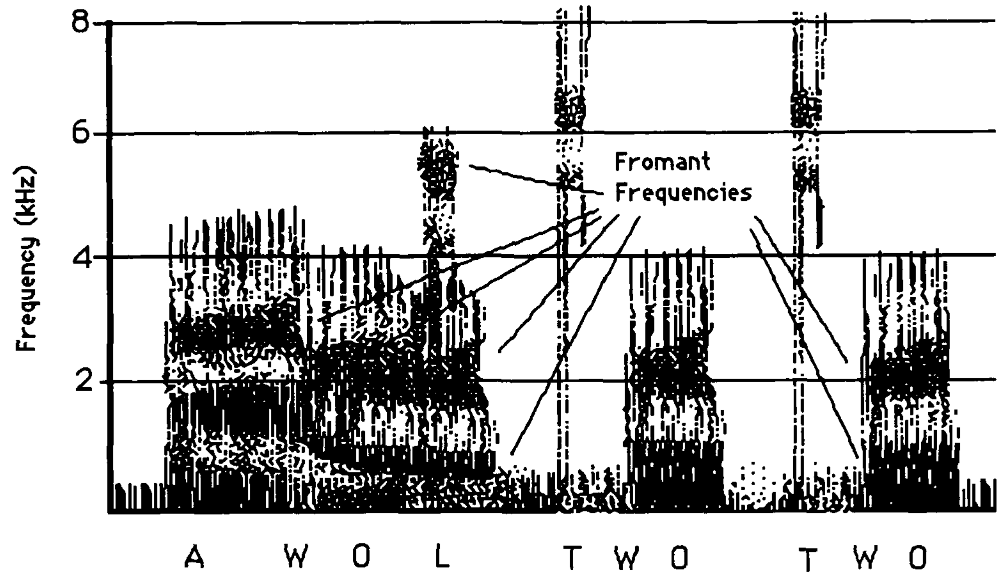

Figure 8. Spectrogram for callsign "AWOL two two".

recognition unit.

To aid in the categorization of vocal sounds, speech may be divided into vocoids and contoids (Parsons, p. 86). Vocoids are vowel sounds which are produced by the vibration of the vocal cords and the creation of resonants within the unobstructed throat and nasal cavities. The resonants establish formant frequencies which can be used to identify particular vowel sounds. A chart of the first three formant frequencies for selected vowels of men and women appears in table two (Parsons, p. 104).

Contoids are consonant sounds which are produced by obstructing the vocal tract in one way or the other. These sounds may be characterized by the location of the constriction, the degree of constriction and manner of release, and whether accompanied by vocal cord vibration (phonation). By using these characterizations, five contoidal attributes may be described (Parsons, p. 88):

- voiced bilabial stop
- voiced dental plosive
- unvoiced apicoalveolar affricate
- unvoiced apicodental spirant
- glottal stop

Remember from page sixteen, that the constricted vocal tract has three effects on voice: 1) frication - the vocal tract is constricted at some point resulting in a broadband noise character-

Table 2.

First Three Formant Frequencies for Selected Vowels

	Adult males			Adult females		
Vowel	F_1	F_2	F_3	F_1	F_2	F_3
[i]	255	2330	3000	340	2610	3210
[I]	350	1975	2560	425	2170	2900
[E]	560	1875	2550	690	2015	2815
[æ]	735	1625	2465	950	1955	2900
[a]	760	1065	2550	1085*		2810
[ʌ]	640	1250	2610	750	1300	2610
[ɔ]	610	865	2540	785*		2565
[U]	475	1070	2410	515	1070	2280
[u]	290	940	2180	390	995	2585

* in females F_1 and F_2 blend in some vowels

ized by the point of constriction, 2) compression (stop) - the vocal tract is completely obstructed allowing pressure to build up which when released results in a short noise burst, and 3) vibration - air is forced through a closure other then the vocal cords. Acoustically, fricatives are characterized by a long stretch of wideband noise, stops are characterized by a short period of silence followed by an abrupt release, and vibration is difficult to characterize (Parsons, p. 119).

The acoustical characteristics of vocoids and contoids can provide cues for the vowel and consonant features of the voiced language. Other cues are the acoustic characteristics of nasal sounds and the formant transitions of consonants. Together, these cues can be used to develop computer algorithms to precisely match the incoming acoustical signal to stored voice data. A block diagram of the recognition unit follows:

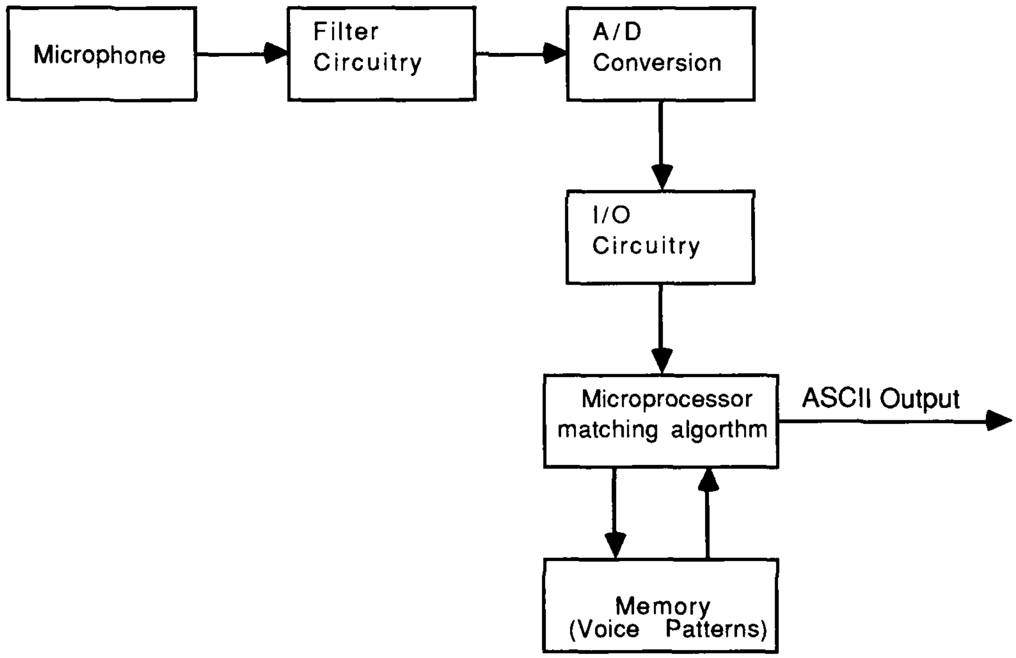

Figure 9. Block diagram of recognition system.

In a continuous speech recognition system, the recognizer must be able to determine the transitions between words. Cues are available for this. Since words are made up of syllables and, in English, every syllable contains a vowel, it would be possible to determine where the syllables are in a sentence by determining where the vowels occurred. Also, word transitions can be identified by a pause, a lengthening of the previous syllable, a momentary glottal closure, or a dip in amplitude (Parsons,p.97). Notice the difference between the voiced word "nitrate" and the phrase "night rate", which are phonetically equivalent. Finally, stress and pitch provide cues for the transition between words, phrases, and sentences.

The Postprocessor

Once the recognizer has identified a voice pattern, it must determine what to do with it. In most cases, it will associate the *voice pattern* word with its ASCII equivalent for presentation somewhere in the machine voice recognition system. A recognition system which allows the user the flexibility to define how this presentation will occur, is very desirable. Usually this definition function is accomplished through some type of transition file. The user specifies what the recognizer will output for a specific voice pattern match.

If the application of the machine voice recognition system is to present voice data, then the ASCII equivalent words may be sent to a postprocessor for presentation on a screen or for transmission to another location. However, if the application of the system is to manipulate the voiced input in some way, then the transition file may be used to tailor the recognition output to postprocessing input requirements.

Some postprocessing applications are:

- presentation of voiced data to one or many users.
- voiced input for forms and data entry.
- storage of voice messages and communications.
- manipulation of voice data for comparison to expert systems or simple procedure rules.
- a switching mechanism to aid in controlling machinery.

These are just a few of the many applications for which the machine voice recognition system may be used.

The use of voice as the input medium to a device provides many advantages, some of which were discussed in Chapter One. Voice can provide a second modality in a visually saturated environment. It requires no use of the hands or special apparatus. It can provide immediate information flow from a user. But, before voice can be put to use, a specific grammar must be developed for the application. This is the subject of the next chapter. With an understanding of how a machine may parse a voiced sentence, we now turn the medium of voice meaning - grammar.

Chapter Three

THE AIR TRAFFIC CONTROL LANGUAGE

Machine voice recognition systems parse verbal language through a voice pattern matching scheme. Although these schemes vary, they all require the user to establish some type of verbal grammar. On simple systems, the grammar may be a one-word command to aid in keyboard entry. An example would be a system which would recognize DOS commands, such as "directory" or "copy", and input their ASCII equivalents, "dir" or "copy" respectively, into the keyboard buffer and onto the screen.

More complicated applications will require the user to carefully select the content and structure of the recognition grammar. Content is important because of the limited number of words a recognizer can ultimately recognize. Grammar structure is important because the structure of the user's language will determine the amount of processing time required to parse a particular sentence.

To aid in the establishment of a verbal grammar, the following steps should be followed:

1) Study the language through reference documents, for command languages, and/or through transcripts, for conversational languages.

2) Develop lists of commonly used sentence structures.
3) Diagram common sentence structures using a verb-centered semantic approach.
4) Use the sentence diagrams to develop your verbal grammar.
5) Implement the verbal grammar under optimal conditions to determine if processing time is acceptable.
6) If processing time is unacceptable, revise the verbal grammar otherwise test the grammar in the actual environment.

These steps will lead to an effective and efficient verbal grammar. Also, by analyzing the language though a verb-centered semantic approach, you define the contextual constraints of the language. The language contextual information will aid in developing artificial intelligence postprocessing applications.

Williams Air Traffic Control Facility

Before dealing with the Air Traffic Control language, it will be helpful to discuss the facility under which tests were conducted and specifically to discuss the duties of the three controller positions in the Williams AFB tower complex (figure ten).

The Williams AFB tower complex is manned by three controllers who communicate with all air traffic passing through the Williams terminal control area. These controller positions are local, ground, and clearance delivery.

The local controller is responsible for clearing aircraft for arrival to and departure from the active runway. For arriving aircraft, the controller begins to talk to aircraft at position points between ten and fifteen miles from the base. He or she normally will ask aircraft to proceed to and report at a position approximately five miles from the runway. At this point, the controller will clear the aircraft to land or proceed to another position, and ensures that the pilot reports the aircraft gear down. The controller provides the pilot with wind and altimeter data. For departing aircraft, the controller clears the aircraft onto the active runway, and provides last minute departure instructions, current winds, and altimeter data.

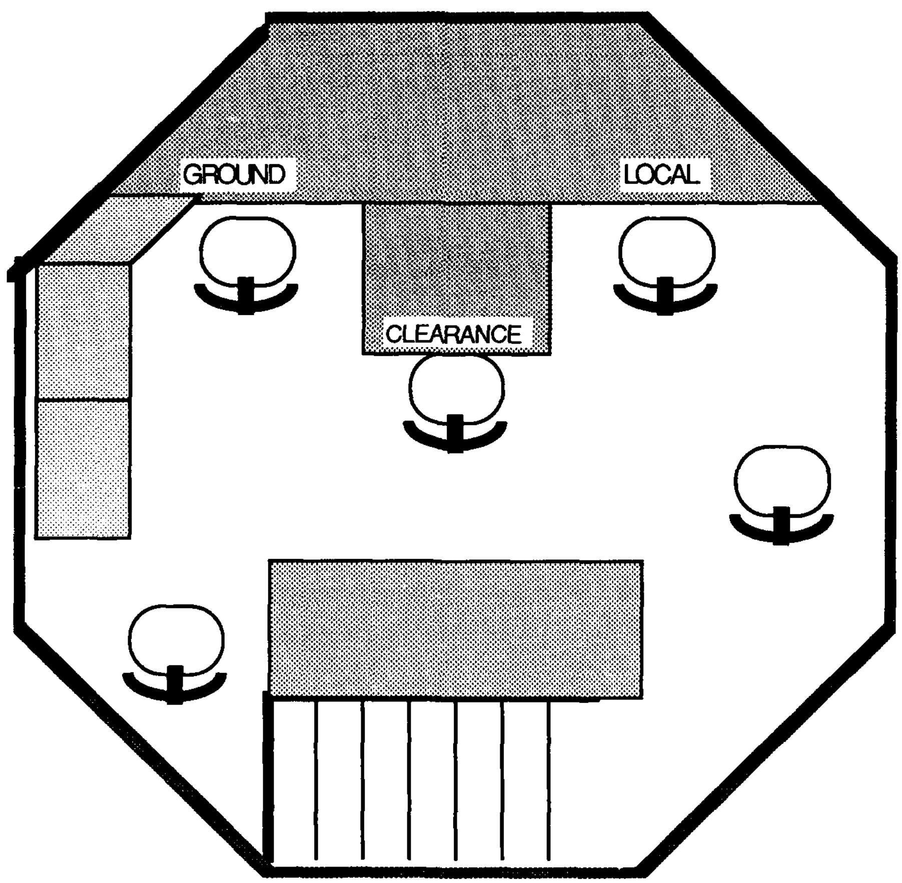

Figure 10. Williams Air Force Base, Air Traffic Control tower facility.

The ground controller is responsible for coordinating all aircraft taxiing from the active runway to parking or parking to the active runway. The controller provides aircraft with taxi clearance, and initial wind and altimeter data. He or she also provides coordination between taxiing aircraft and any motor vehicles which may be driving on the taxiways or parking area. An example would be coordination between aircraft and responding fire equipment.

The clearance delivery controller is responsible for the coordination of all air traffic control clearances from the base. He or she normally talks only to aircraft departing the local area. The controller issues a clearance to a flight which includes the destination base, departure route, cleared altitudes, the radio frequency for departure control, and the numerical settings for the aircraft transponder (normally called a "squawk"). Local flights receive an abbreviated clearance which does not require contact with the clearance delivery controller.

Language Study and Sentence Structure

The Air Traffic Control language was studied through the use of handbooks and controller transcripts. Federal Aviation Administration Handbook 7110.65 details controller sentence structure.

Although handbooks prescribe what a controller is to say, shortcuts and local procedures can effect commonly used sentence structure. For this reason, nine hours of controller transcripts were studied. Transcript times were selected which would present the most varied sentence structure. The selected time frame was on Friday afternoons, when transient and cross-country flights peaked.

Transcripts for each controller position appear in Appendix B. This information was used in diagramming the common sentence structures using a verb-centered semantic method.

Verb-Centered Sentence Diagramming

The use of verb-centered semantics in determining the structure of a sentence requires the understanding of *case* and how it is used to gain information from a sentence. By *case,* is meant the pattern of inflection of nouns, pronouns, and adjectives to express different meaning. In English, we tend to be unfamiliar with the use of case because our language has only three *visible* cases for pronouns and one *visible* case for nouns. The English cases include:

case	pronoun	noun
nominative case	he (He gave it to me.)	
objective case	him (I offered him the gift.)	
possessive case	his (Now, it is his.)	's (It is Rob's.)

Whatever method we choose to represent the structure of a sentence, it must be able to consistently encode the meaning of a sentence. It is not possible to do this by using a scheme which identifies the sentence subject and object. For instance:

- John opened the door.
- The door was opened by John.

The subject is "John" in the first sentence, while the subject is "door" in the second sentence. Yet, the two sentences convey the same meaning. Therefore, another method must be found in order to consistently portray the meaning of a sentence. Case provides us with a method to represent the underlying meaning of every sentence in a determinative manner.

Charles Fillmore is accredited with establishing the modern theory of case. In his 1968 article "The Case for Case", he identifies case as: "a set of universal, presumably innate, concepts which identify certain types of judgments about such matters as who did it, who it happened to, and what got changed." He goes on to identify eight cases:

- **Agentive** (agnt): instigator of action identified by verb.
- **Instrumental** (inst): inanimate force or object causally involved in the action or state identified by the verb.
- **Dative** (dat): animate being affected by state or action identified by verb.
- **Factitive** (fact): object or being resulting from the action or state identified by the verb.
- **Locative** (loc): location of state or action identified by the verb.
- **Objective** (obj): thing representable by a noun whose role in the action or state is identified by the verb.
- **Benefactor** (bene): the animate which benefits from agentive action.
- **C(omitative)** (com): a subordinate noun phrase with the same case.

Fillmore established a set of formulae for describing a sentence and its parts. A sentence is made up of a modality and a proposition (S = M + P). The modality part of the formula represents such concepts as negation, tense or mood. The expansion of the proposition part is represented by the sentence verb and one or more cases of noun phrases, such as:

$$P = V + Case_1 + Case_2 + \ldots + Case_n,$$
$$P = V + A,$$
$$P = V + O + A,$$
$$P = V + D,$$
$$P = V + O + I + A,$$

where V is a verb and A, O, D, I, are the first letters of a particular case.

In this way, the meaning and structure of a sentence may be uniquely described by identifying the modality, verb and all case relations of the sentence noun phrases. Fillmore says, "the sentence in its basic structure consists of a verb and one or more noun phrases, each associated with the verb in a particular case relationship." Thus, the verb becomes central to our investigation of the meaning of a sentence.

For instance, the sentence "Awol 44, Willy Ground, taxi to runway 30 center" could be represented as:

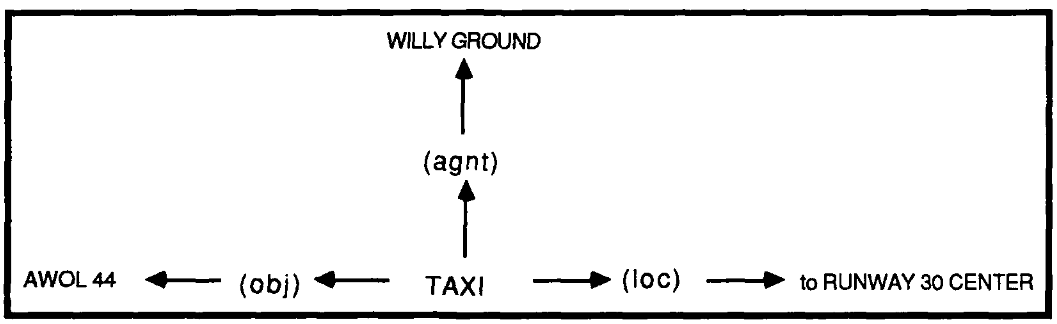

Figure 11. Sentence diagram for "AWOL 44, willy ground, taxi to runway 30 center."

This representation of the sentence structure provides a clear sketch of its meaning.

The sentence structures found in the Air Traffic Control environment were complex and lengthy. This was due to the fact that single sentences were devised to relay a large amount of information. The result is compound sentences with many attributes. To model these sentences Fillmore type cases with the addition of "attribute" (attr) and "link" (link) cases were used.

The attribute case categorizes the optional noun phrases found in many ATC sentences such as altimeter settings, wind conditions, and bird condition. By optional it is meant that these noun phrases do not occur in every utterance of that particular verb sentence due to certain constraints. For instance, the bird condition noun phrase is not spoken unless birds are seen in the area and their presence poses a possible safety hazard.

The link case is very similar to the comitative Fillmore case in that it serves as a link from the noun phrase in one case to more specific information. Thus, links are used to indicate more specific detail as seen in many position or route phrases in the Air Traffic Control environment.

The additional cases and their multiple use in a sentence is a pragmatic solution to modeling complex voice communications. It provides us with a means to visualize the sentence structure so that a recognition grammar can be constructed. Linguists cannot agree on what cases are necessary in order to model the English language. Some authors develop hierarchical sets of cases with supercases and subcases (Somers, 1987, p.113). Somers comments: (p.119)

> "... The important point is that the notion of supercase-subcase merely reflects the point that Case is altogether a question of making significant generalisations, ... The aim must be to arrive at a level of abstraction and generalisation that serves the purpose to which the model is to be put."

The formal linguistic approach to modeling language is concerned with the ability to be able to represent the meaning of any sentence no matter what the order of the particular sentence noun phrases. For instance, in the sentences:

- John opened the door.
- The door was opened by John.

These sentences have the same meaning, but their noun phrases are ordered differently. Rules have been developed to deal with this phenomenon.

However, the development of a grammar definition program requires that a strict language order be established. In the case of an Air Traffic Control language this order is already established and adhered to. In order to indicate this order in sentence diagrams, a clockwise noun phrase order convention was used. The reader should start reading the diagram on the left and proceed in a clockwise fashion around the verb. The "attr" relation specifies optional noun phrases and the "link" relation specifies additional detail. Note that to allow for the generation of a wide variation of sentences, most noun phrases are optional. Incidentally, the placement order of the verb is not indicated in a diagram, so care must be used when placing the verb in the grammar definition. In this application, the end result was the creation of sentence structure diagrams and a grammar definition program for the three controller positions at the Williams Air Traffic Control facility.

LOCAL CONTROLLER CONTEXTS

Local Controller Sentence Diagrams

The sentence diagrams for the local controller are presented in the next few pages. Essentially, the language which the local controller uses to communicate with aircraft can be defined within the context of four verbs: *cleared*, *report*, *taxi*, and *hold*. The context of each verb is graphically portrayed in figures twelve to fifteen. Following each figure, typical sentence structures are presented.

Local controller 'cleared' context:

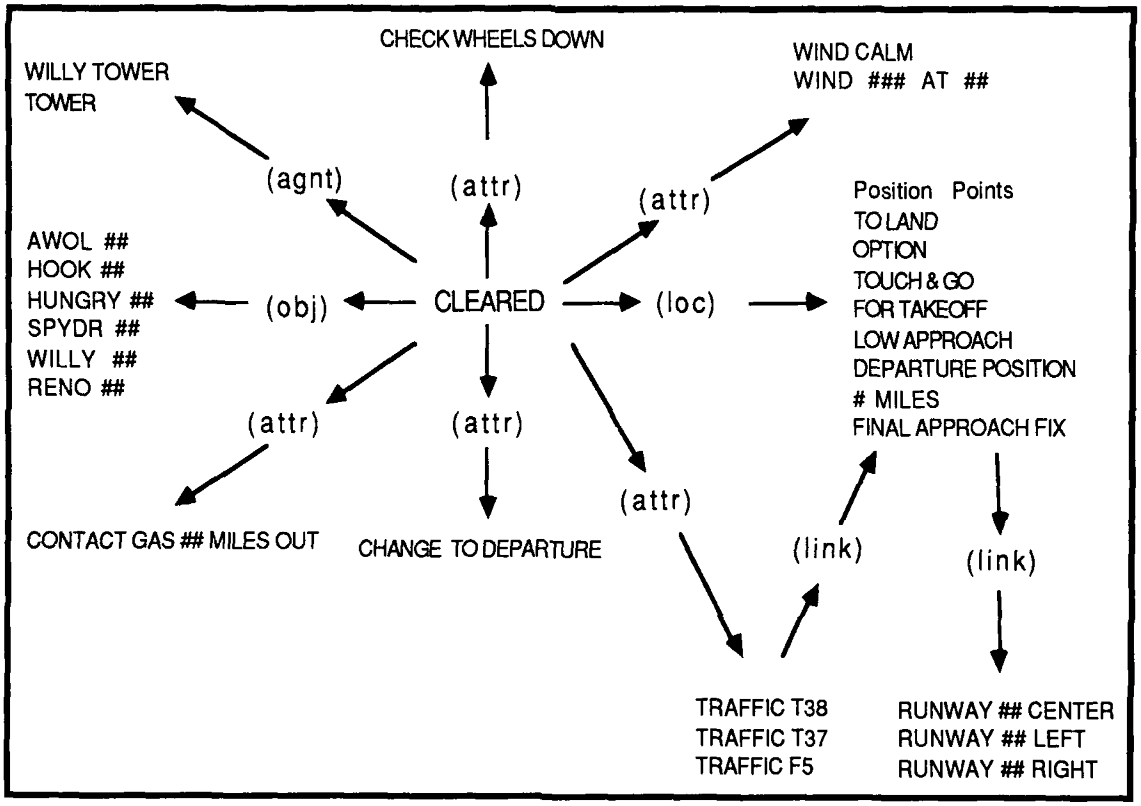

Figure 12. Sentence diagram for 'cleared' context (# = digit).

Typical sentence structures:

- AWOL 44, Willy tower, wind calm, cleared for takeoff, change to departure.
- AWOL 42, wind calm, cleared for touch-and-go runway 30 center, traffic T-38 departure position.
- AWOL 35, Willy tower, cleared for takeoff.
- AWOL 43, check wheels down, wind calm, cleared for the option runway 30 center.
- Willy 80, wind 140 at 6, cleared to land runway 30 center.

Local controller 'report' context:

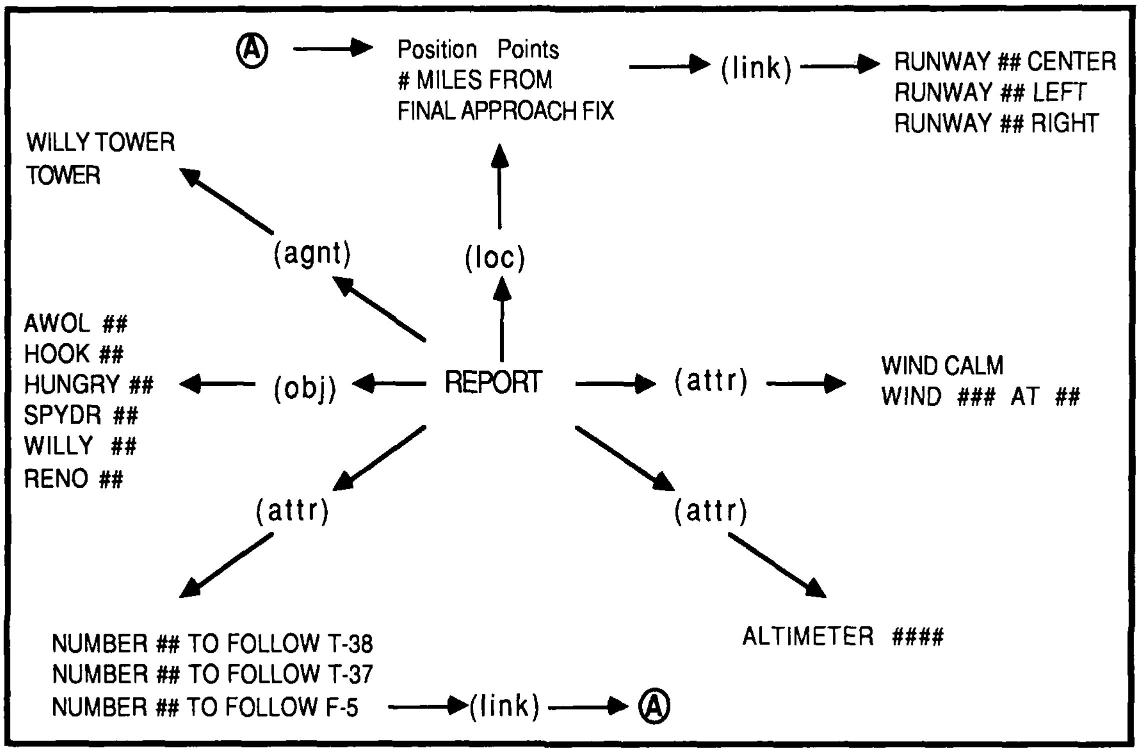

Figure 13. Sentence diagram for 'report' context (# = DIGIT).

Typical sentence structures:

- AWOL 44, Willy tower, report final approach fix runway 30 center, wind calm, altimeter 3025.
- RENO 36, Willy tower, report final approach fix runway 30 center, wind 120 at 13, altimeter 2992, number 2 to follow a T-38 five miles.
- AWOL 17, tower, report 5 miles fron runway 30 center, wind calm, altimeter 3025, number 2 to follow T-38 final approach fix.
- Hungry 32, Willy tower, report final approach fix.

Local controller 'taxi' context:

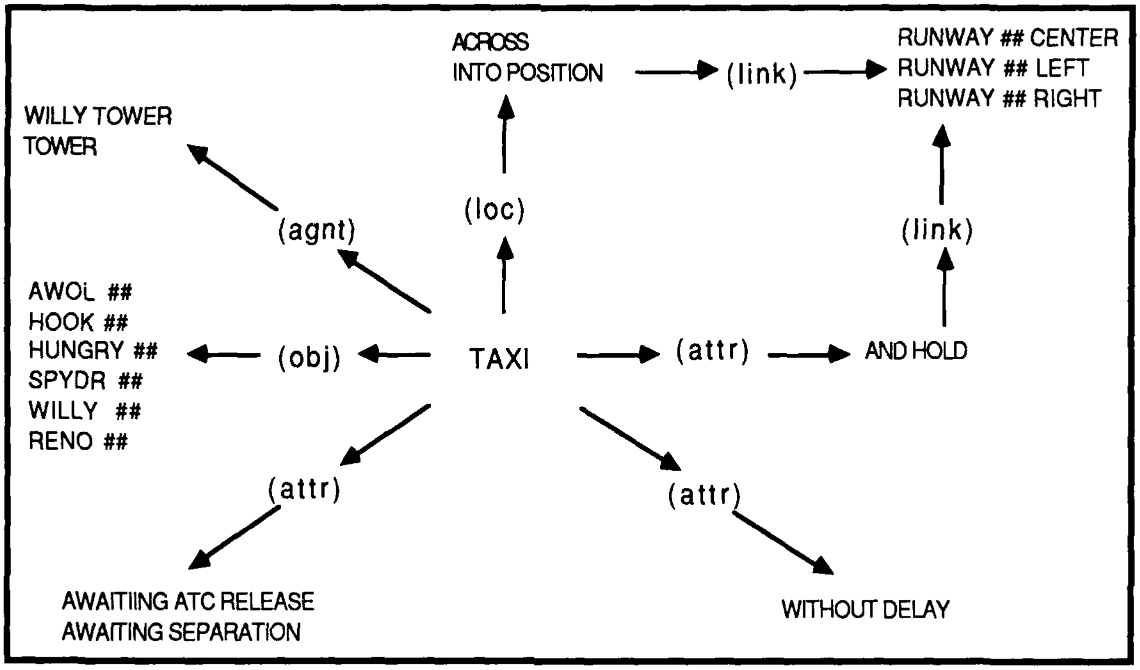

Figure 14. Sentence diagram for 'taxi' context (# = DIGIT).

Typical sentence structures:

- AWOL 64, tower, taxi across runway 30 center, without delay.
- AWOL 75, Willy tower, taxi across runway 30 center.
- Hungry 01, Willy tower, taxi into position and hold runway 30 center, awaiting ATC release.
- AWOL 63, tower, taxi into position and hold 30 center.
- Reno 33, Willy tower, taxi into position and hold, awaiting separation.

Local controller 'hold' context:

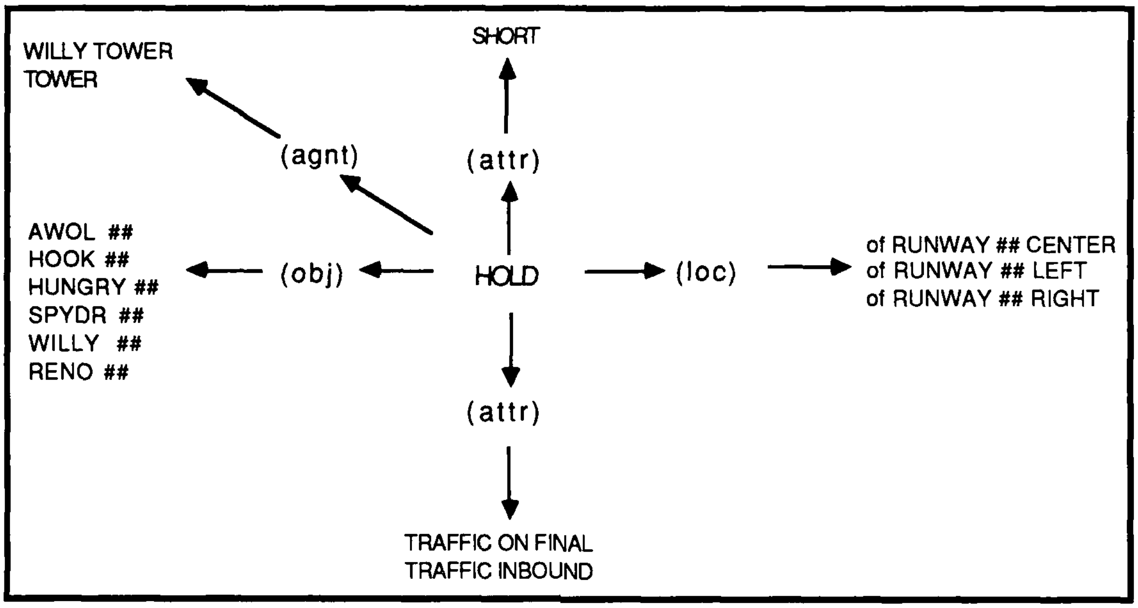

Figure 15. Sentence diagram for 'hold' context (# = DIGIT).

Typical sentence structures:

- Willy 80, Willy tower, hold short.
- Willy 51, Willy tower, hold short, traffic on final.
- AWOL 66, Willy tower, hold short, traffic inbound.
- AWOL 75, tower, hold short.
- Reno 22, Willy tower, hold short of runway 30 left.

Local Controller Grammar Definition Program

The development of the grammar definition program is usually dependent on a proprietary language provided by the recognizer manufacturer. The language 'programs' the recognition unit to accept words within the context defined by the user. It also aids in ensuring that the capabilities of the recognizer are not exceeded by the number of vocabulary words or their complexity.

The advantage of using a sentence diagraming technique to identify the language context is found in the ease of converting the diagrams into a proprietary programing language. The program is concerned with identifying the flow of the user's sentences and indicating which words are optional or mandatory within this flow. The sentence diagrams pictorially display this information through the use of a clockwise order notation and the case relations.

In the Air Traffic Control application, the sentence diagrams were used to generate a grammar definition program using the Voice Planner™ software provided by Vocalink®. The sentence diagrams were converted into a text file using Interstate Voice Products Standard Notation (IVPSN). The text file is converted into a voice recognition program for the Vocalink® S4000 by the Voice Planner™ software. A portion of the IVPSN text file used to create the local controller recognition grammar program is presented on the next page:

```
;
;   Grammar Definition File for local controller voice recognition
;
.OBJ/ .DIGIT@2 .AGNT/ check-wheels-down/ taxi/ report/ cleared/ .LOC/ .HLD/
 & without-delay/ .WIND/ .ALT/ .TRF/ change-to-departure/

... (see Appendix C for entire program)

.LOC=
     .DIGIT@2 .DIR/
       into-position/ across/ runway .DIGIT@2 .DIR/
     .DIGIT@1 miles .DIGIT@2 .DIR
       final-app-fix
     to-land
     option
     touch-and-go

.DIR=
     center
     left
     right

.WIND=
     wind calm
     wind .DIGIT@3 at .DIGIT@1,2

.ALT=
     altimeter .DIGIT@4

.DIGIT=
     zero
     one
     two
     three
     four
     five
     six
     seven
     eight
     nine
```

The bold type in the program indicates the main portion of the program which defines the order and extent of the local controller's language. It contains the verbs, some noun phrases, and many structure names which are specified later in the program (the verb 'hold' is contained in the structure '.HLD' which is not shown, see Appendix C). The slash after a word or structure name indicates that the utterance is optional. Note that many structures contain or call other structures.

GROUND CONTROLLER CONTEXTS

Ground Controller Sentence Diagrams

The sentence diagrams for the ground controller are presented in the next few pages. This language can be defined within the context of three verbs: *taxi*, *proceed*, and *approved*. The context of each verb is graphically portrayed in figures sixteen to eighteen. Following each figure, typical sentence structures are presented.

Ground controller 'taxi' context:

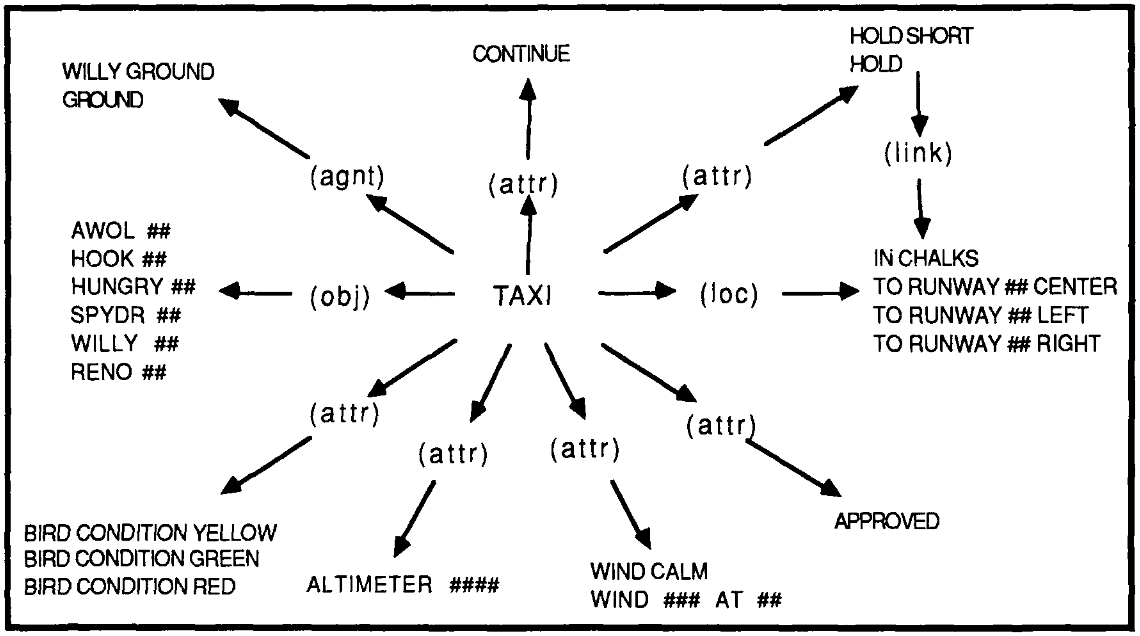

Figure 16. Sentence diagram for 'taxi' context (# = DIGIT).

Typical sentences structures:

- AWOL 50, Willy ground, taxi to runway 30 center, wind calm, altimeter 3023.
- Reno 06, ground, continue taxi to runway 30 center, wind calm, altimeter 3024.
- Willy 91, ground, taxi approved.
- Hook 64, Willy ground, taxi to runway 30 left, winds 270 at 4, altimeter 2982, bird condition yellow.
- Hook 65, ground, taxi, hold-short runway 30 left.

Ground controller 'proceed' context:

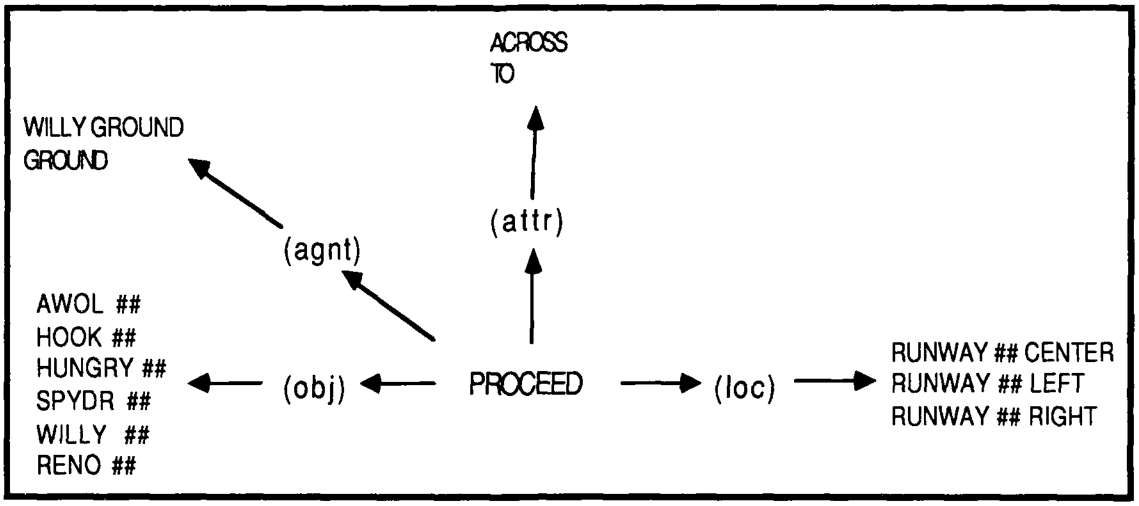

Figure 17. Sentence diagram for 'proceed' context (# = DIGIT).

Typical sentence structures:

- AWOL 22, Willy ground, proceed across runway 30 left.
- Reno 53, ground, proceed across runway 30 left.
- Willy 14, ground, proceed across runway 12 right.

Ground controller 'approved' context:

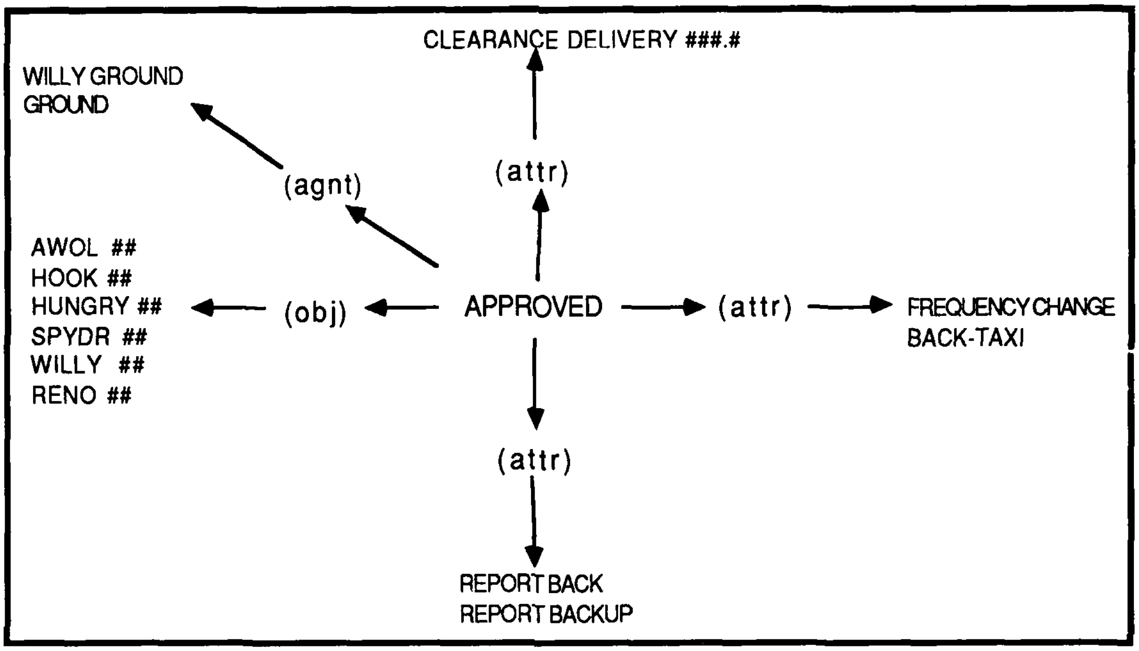

Figure 18. Sentence diagram for 'approved' context (# = DIGIT).

Typical sentence structures:

- AWOL 56, Willy ground, frequency change approved.
- Hungry 44, ground, back-taxi approved.
- Willy 91, ground, clearance delivery 289.4, frequency change approved, report backup.
- Spydr 22, Willy ground, frequency change approved, report back.

Ground Controller Grammar Definition Program

As with the local controller grammar definition program, the ground controller grammar was established using the VocaLink® proprietary software. A portion of the text file is presented on the next page with the entire file appearing in Appendix C.

```
;
;   Grammar Definition File for ground controller voice recognition
;
.OBJ/  .DIGIT@2  .AGNT/  continue/  taxi/  proceed/ .APP/ .HLD/ .LOC/ .WIND/
&  .ALT/  .BIRD/  .RPT/  .CONT/

... (see Appendix C for entire program)

.OBJ=
     awol
     hook
     hungry
     spydr
     willy
     reno

.AGNT=
     ground
     willy ground

.APP=
       frequency-change/ back-taxi/ approved

.HLD=
     hold
     hold-short

.LOC=
     .DIGIT@2 .DIR/
     across/ to/ runway .DIGIT@2 .DIR/

.DIR=
     center
     left
     right
```

CLEARANCE DELIVERY CONTEXTS

Clearance Delivery Sentence Diagrams

The sentence diagrams for the clearance delivery controller are presented in figures nineteen and 20. The language can be defined within the context of the verb: *cleared*, and the noun: *clearance*.

The use of a noun in defining context is unusual, but is used here because it easily describes several sentence structures. Note that the 'clearance' noun phrase comes directly from the 'cleared' verb and could probably be included in that context. However, the cleared context is already complex and further expansion was undesirable. Following each diagram, typical sentence structures are presented.

Clearance delivery 'clearance' context:

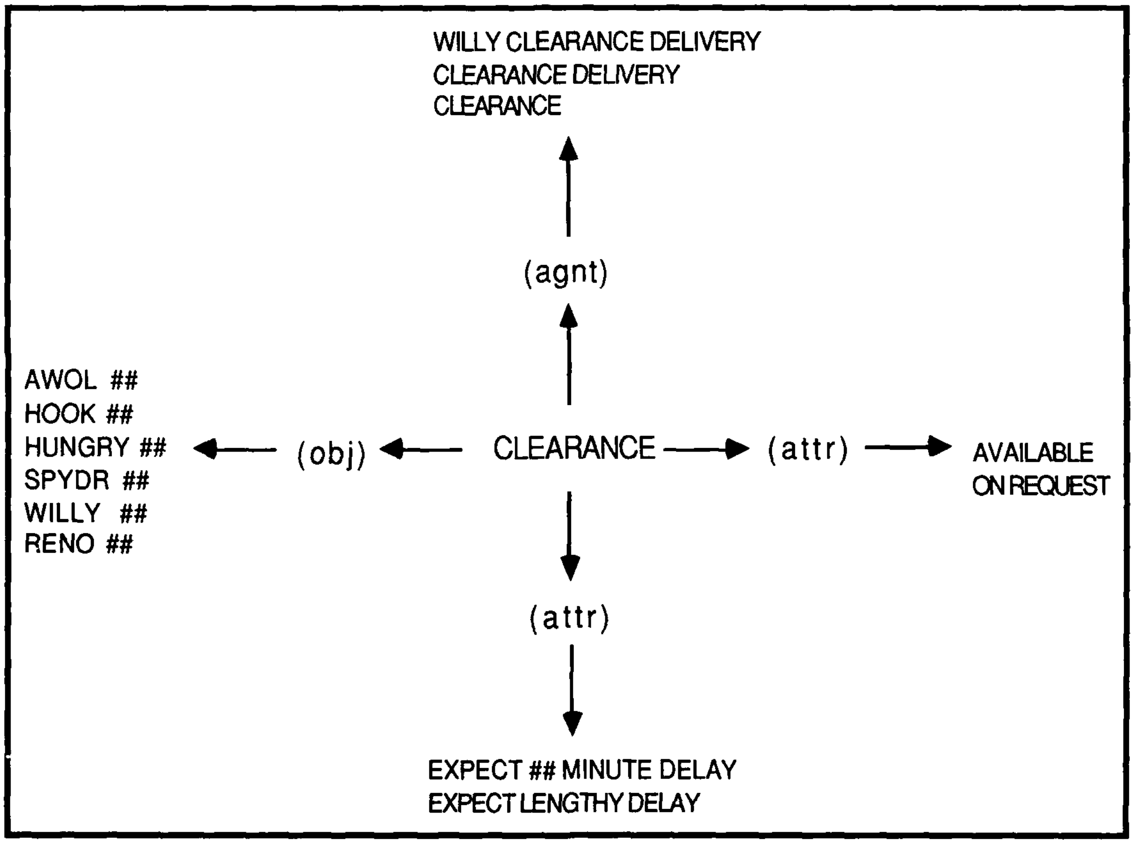

Figure 19. Sentence diagram for 'clearance' context (# = DIGIT).

Typical sentence structures:

- Hook 22, Willy clearance delivery, clearance on request.
- Hook 22, clearance delivery, clearance available.
- Hungry 04, clearance delivery, clearance on request, expect 30 minute delay.

Clearance delivery 'cleared' context:

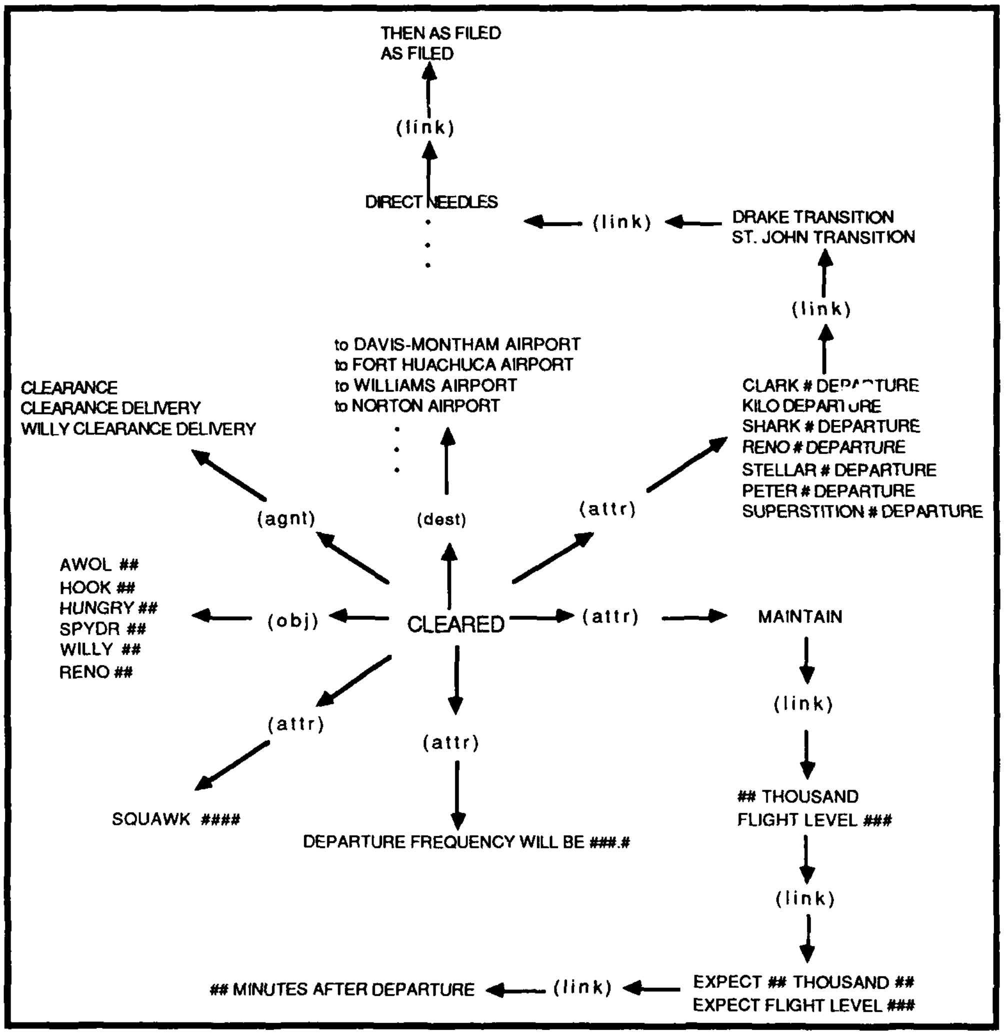

Figure 20. Sentence diagram for 'cleared' context (# = DIGIT).

Typical sentence structures:

- Willy 91, Willy clearance delivery, you're cleared to Williams airport, shark 7 departure, maintain flight level 190, departure frequency will be 317.5, squawk 4230.
- Willy 07, Willy clearance delivery, you're cleared to Norton airport, Stellar 5 departure then as filed, maintain flight level 190, expect flight level 370, squawk 0701.
- Reno 58, you're cleared to Reese airport, Superstition 9 departure, Saint John transition, then as filed, maintain flight level 270, expect flight level 330, 10 minutes after departure, departure frequency will be 317.5, squawk 4120.

Clearance Delivery Grammar Definition Program

The clearance delivery grammar was established using the VocaLink® proprietary software. A portion of the text file is presented below with the entire file appearing in Appendix C.

```
;
;   Grammar Definition File for clearance delivery voice recognition
;
.OBJ/  .DIGIT@2  .AGNT/  cleared/  clearance/  .DEST/  .DEP/  .TRANS/  .DCT/
& .FILE/  .MTN/  .EXP/  .TIME/  .FREQ/  .SQW/
roger

 ... (see Appendix C for entire program)

.OBJ=
      awol
      hook
       hungry
      spydr
       willy
      reno

.AGNT=
      clearance
        clearance-delivery
         willy clearance-delivery

.DEST=
        davis-montham airport
         fort-huachuca airport
         williams airport
        norton airport

.DEP=
      .NAME .DIGIT@1 departure

.NAME=
       clark
       kilo
       shark
      reno
        stellar
      peter
        superstition
```

With a grammar defined for the controller positions, the next step is to test the grammar in an optimal environment and with successful results in an operational environment. Also, determination of the specific output of the recognizer is necessary. These topics are discussed in the next chapter.

Chapter Four

This chapter describes the method used to test the feasibility of integrating a voice recognition system into the Air Traffic Control environment. The ground controller grammar discussed in Chapter Three was refined and implemented on the VocaLink® S4000 voice recognition unit. Over a period of two months, a controller at the Williams AFB tower complex trained on the unit and tested the operation of the grammar. The results indicated that the recognition system would be highly reliable for the parsing of single words and short phrases. The parsing reliability of long phrases was near 100%, however, the unit accepted less then 60% of the controller's long phrases when spoken in a operational setting.

CONTROLLER GRAMMAR

The ground controller grammar described in Chapter Three was formed by a reiterative process of training and testing. Initially, the grammar was developed using Interstate Voice Products' Voice Planner™ software and a text file editor. Training was accomplished on the unit in an office environment and the grammar was tested to see how well it performed. After several training and testing passes, the ground controller grammar was established.

The Voice Planner™ software transformed each of the grammar definition files developed in Chapter Three into binary files which the voice recognizer uses. In this process, the program calculated a complexity percentage index indicating the difficulty the unit will have processing the grammar. The complexity percentage indexes for the three grammars were:

Local controller	107%
Ground controller	114%
Clearance delivery	126%

The VocaLink® software documentation specified that grammars with complexity indexes between 80 and 120% needed to be tested to determine if the recognition unit would work reliably. It was felt to be possible to improve on the ground controller complexity index by eliminating vocabulary which was used very little. The result was the grammar definition which appears below:

```
;
;  Modified Grammar Definition File for ground voice recognition
;
.OBJ/ .DIGIT@2 .AGNT/ continue/ taxi/ .APP/ .HLD/ .LOC/ .WIND/ .ALT/
roger
runway change

.OBJ=
      awol
      hook
      hungry
      spydr
      willy
      reno

.AGNT=
      ground
      willy ground
.APP=
        frequency-change/ back-taxi/ approved
```

```
.HLD=
      hold
      hold-short

.LOC=
      .DIGIT@2 .DIR/
      runway .DIGIT@2 .DIR/

.DIR=
      center
      left
      right

.WIND=
      wind calm
      wind .DIGIT@3 at .DIGIT@1,2

.ALT=
      altimeter .DIGIT@4

.DIGIT=
      zero
      one
      two
      three
      four
      five
      six
      seven
      eight
      nine
```

This grammar has a complexity index of 82% and was extensively tested in an Air Traffic Control operational environment.

Once the grammar definition file was established, it was used in conjunction with a translation file (see Chapter Five) to create a MASTER cartridge. The MASTER cartridge is used by the recognition unit during user training to determine the vocabulary voice patterns it needs to parse the grammar. During this user training, a USER cartridge is created which contains the voice patterns of the user in a battery refreshed memory.

The storing of the user's voice patterns on a cartridge is a very convenient measure. *In the case of multiple users,* each one could have his or her own cartridge. In the Air Traffic Control application, the controllers would insert their USER cartridge into the recognition unit (see figure 21) before starting duty or when switching controller positions.

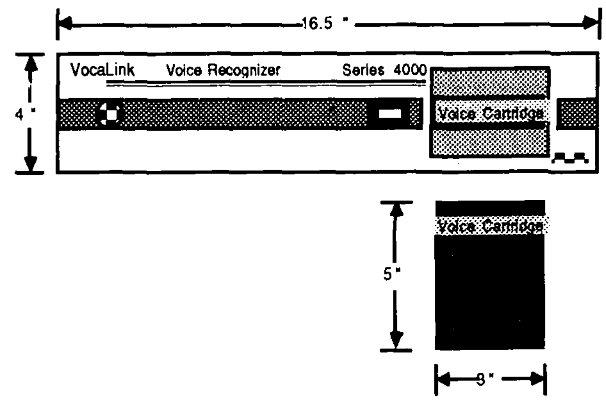

Figure 21. Voice cartridge.

CONTROLLER TRAINING

Equipment used in the training and testing phase of this study included: a VocaLink® S4000 Voice Recognizer, a Standard 286 - IBM AT compatible computer with monitor, and a Macintosh SE computer. The Standard 286 was used with the Voice Planner™ software as a training terminal and the Standard 286 was also used as a host computer when it ran the PROLOG language. At other times, the Macintosh SE computer was used as a host computer to capture and record the S4000's output (see figure 22).

A volunteer fully-qualified controller was used as the subject during the training and testing phase. He made his first training pass with the training terminal in an office environment. Interstate Voice Products recommends an office environment for this first pass so that the recognition unit can obtain a good set of voice

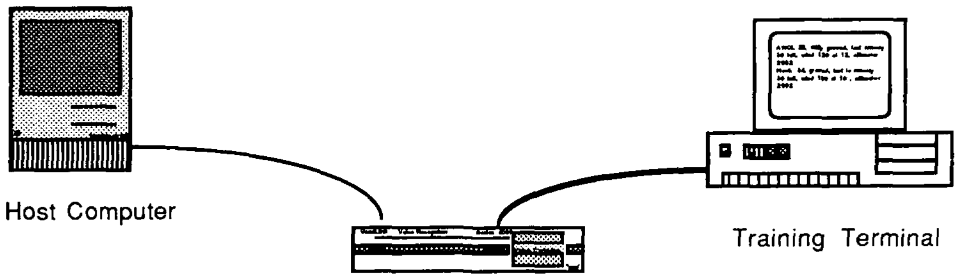

Figure 22. Test equipment.

patterns without background noise. The first pass required the controller to voice all vocabulary words into the Shure® VR230A headset microphone at least twice. Next the controller voiced a series of phrases prompted by the training terminal as defined by the ground controller voice grammar. All training passes lasted approximately 30 minutes.

During the second training pass, the controller trained in the tower facility at Williams AFB. The second pass in the operational environment allows the recognition unit to establish an expected background noise. In all, a total of seven training passes were accomplished with the ground controller grammar.

The tower facility environment is characterized by random levels of noise with the strongest source being the voice of other controllers. Noise is emitted primarily from five categories:

• other controllers	max of 76 db - B scale.
• air conditioning system	max of 67 db - B scale.
• telephones	max of 72 db - B scale.
• teletype	max of 72 db - B scale.
• weather strip writer	max of 73 db - B scale.

Other noises do occur, such as the audible alert warning on radio navigation aids, but these are generally well below the noise level of more frequent noises.

The VocaLink® S4000 recognition unit is advertised to be operational in environments with up to 85 db background noise. Therefore, the tower noise sources were not considered to be detrimental to the operation of the unit. It should be noted that the tower noises were very random and thus must have been difficult for the recognition unit to isolate.

The testing of a voice recognition unit can be very difficult. Ultimately, the evaluator most decide on what criteria the unit will be evaluated and, for continuous recognition systems, what type of voiced data is to be considered. The following criteria were used and are recommended: recognition accuracy, recognition acceptance, and false recognition. Also, the following types of voiced data were used and are recommended: single words, short phrases, and long phrases.

Recognition accuracy is the percentage of correct word matches given the acceptance of a particular phrase. From this statement it is apparent that not all phrases are accepted by the unit. Recognition acceptance is the percentage of correctly voiced (within the confines of the grammar) phrases or words accepted as input by the recognition unit. False recognition is the percentage of words or phrases accepted as input by the unit but not part of the defined grammar. False recognitions occur when the unit picks up casual conversation by the controller and outputs a false word or phrase as if the controller were talking to an aircraft.

The differing categories of voice data are necessary in continuous speech systems because it is obviously easier for the unit to process single words or short phrases. The VocaLink® S4000 was tested with single words, short phrases and long phrases.

VOICE RECOGNIZER TESTING

The first series of tests were conducted after the fifth training pass. The tests consisted of having the controller speak predetermined words, short phrases, and long phrases while in a controlled operational environment. i.e. not to actual aircraft. The test for short and long phrases was captured by the Macintosh SE as host computer and is presented below:

Recognition test ground controller - long phrase. At 5 passes in training

awol 2 4 willy ground taxi runway 3 7 center wind 1 2 0 at 3 altimeter 3 0 2 3
awol 5 3 willy ground taxi runway 3 0 center wind 1 2 0 at 3 altimeter 3 0 2 3
awol 8 7 willy ground taxi runway 3 0 center wind 1 2 0 at 3 altimeter 2 0 2 3
awol 4 3 willy ground taxi runway 3 7 center wind 1 2 0 at 3 altimeter 3 0 2 3
awol 5 7 willy ground taxi runway 3 0 center wind 1 2 0 at 3 altimeter 3 0 2 3
awol 8 1 willy ground taxi runway 3 0 center wind 1 2 0 at 3 altimeter 3 0 2 3
hook 1 7 willy ground taxi runway 3 7 center wind 1 2 0 at 3 altimeter 3 0 2 3
hook 4 9 willy ground taxi runway 3 0 center wind 1 2 0 at 3 altimeter 3 0 2 3
hook 9 2 willy ground taxi runway 3 7 center wind 1 2 0 at 3 altimeter 3 0 2 3
hook 3 1 willy ground taxi runway 3 0 center wind 1 2 0 at 3 altimeter 3 0 2 3
hook 2 1 willy ground taxi runway 3 0 center wind 1 2 0 at 3 altimeter 3 0 2 3
hook 8 3 willy ground taxi runway 3 0 center wind 1 2 0 at 3 altimeter 3 0 2 3

100% acceptance - 1945 hrs Mar 28

Recognition test ground controller..... short phrase

awol 2 4 ground taxi approved
awol 5 3 ground taxi approved
awol 8 7 ground taxi approved

awol 4 3 ground taxi approved
awol 5 7 ground taxi approved
awol 8 1 ground taxi approved
hook 1 7 ground taxi approved
hook 4 9 ground back-taxi approved
hook 9 2 ground taxi approved
hook 3 1 ground taxi approved
hook 2 1 ground taxi approved
hook 8 3 ground taxi approved

100% acceptance 1952 hrs 28 Mar

******Italic*** text typed in by operator, plain text is output from S4000**

The phrases used in this test were selected on the basis of what a Williams AFB ground controller speaks most frequently. The same phrase was repeated except for the change in aircraft callsign digit. The results of this test appear in table three.

Table 3.

Results from Controlled Environment Test

	Recognition Accuracy	Recognition Acceptance	False Recognition
Single Word	100%	100%	~
Short Phrase	98.6%	100%	~
Long Phrase	98%	100%	~

*environment did not present false recognition risk

During the test, the recognizer confused the number seven with zero (as in "runway 37") and two with three (as in "altimeter 2023") resulting in a slight decrease in recognition accuracy. The unit accepted every correctly voiced phrase. It is quite apparent that the unit is capable of parsing complex sentences with many possible options. It should be noted that this test was conducted in a controlled environment, with no interruptions, and little background noise. The controller tended to annunciate his words and speak more slowly then his typical ground to aircraft communication.

Further testing was conducted after the sixth training pass. In this test, the controller acted as if the voice recognizer did not exist and he called aircraft and talked to other controllers normally. The transcript from this test as captured by the Macintosh SE follows:

Ground controller test - 30 Mar '88 0830

willy 1 5 willy ground taxi approved hold-short runway 1 2 right
hook 5 1 ground
hook 5 1
willy 5 5
awol 1 3 willy ground taxi approved hold-short runway 1 2 right
hook 4 7 willy ground taxi runway 1 2 right wind 1 4 0 at 6 altimeter 2 9 8 2
awol 7 2 willy ground taxi approved hold-short runway 1 2 right
hook 6 6 willy ground taxi runway 1 2 right wind 1 5 0 at 4 altimeter 2 9 8 2
hook 2 5 willy ground taxi runway 1 2 right wind 1 4 0 at 7 altimeter 2 9 8 2
hook 5 0 willy ground taxi approved
awol 7 6 willy ground taxi approved hold-short runway 1 2 right
hook [illegible] 5 willy ground
hook 5 6 ground taxi approved
awol 1 1 willy ground taxi approved hold-short runway 1 2 right
hook 6 3 willy ground taxi runway 1 2 right wind 1 5 0 at 7 altimeter 2 9 8 2
awol 2 3 willy ground taxi approved hold-short runway 1 2 right
willy 8 2 willy ground taxi approved hold-short runway 1 2 right
hook 6 4 willy ground taxi runway 1 2 right wind 1 2 0 at 4 altimeter 2 9 8 1

end test 1000

******Italic*** **text typed in by operator, plain text is output from S4000**

The results of this test appear in table four.

Table 4.

Operational Environment Test

	Recognition Accuracy	Recognition Acceptance	False Recognition
Single Word	100%	100%	0%
Short Phrase	100%	88.9%	0%
Long Phrase	98.6%	59%	0%

In the operational environment, recognition acceptance degrades, although recognition accuracy is excellent and false recognition did not occur.

The lower recognition acceptance means that several times when the controller spoke a properly voiced long phrase, the voice recognizer did not accept it as input. This was a rather disturbing result and many causes were investigated. First, the controller spoke faster and with less articulation than when he spoke in a controlled environment. Second, background noise was present although not at an unusual level. Third, the controller carried on conversations with other controllers which may have confused the unit. Fourth, the recognizer may require further voice pattern user training.

Attempts were made at correlating lower recognition acceptance with its possible cause. User training did not appear to be a factor because acceptance rates remained approximately equal after the fifth training pass. The headband microphone had an on-off switch which allowed the controller to turn the mike off during casual conversation. No noticeable effect was found between leaving the mike on or having the controller switch off the mike during casual conversation. An increased level of background noise did not correlate with lower recognition acceptance. However, the factor which did have an effect on recognition acceptance was the controller voice speed.

During the controlled environment test, the controller read the aircraft calls from a script. This resulted in a slower voice speed than found in the operational environment. In their operational environment, controllers speak automatically with a phraseology learned from intensive training. They are taught to speak clearly and fast. The result is that the recognizer has a difficult time parsing a controller's phrase when spoken at full speed. For optimum voice recognition results, the controller must slow his speech to a normal conversational rate.

The manner of articulation was also investigated. It was found that it was possible for the controller to over-articulate a phrase resulting in decreased recognition accuracy. Thus, while it is beneficial for the controller to speak at a normal rate with normal articulation, it is possible for the user to over-articulate his or her words. Over-articulation seems to occur when the user emphasizes the ending consonants of words in a phrase. Considering that over-articulation usually occurred when the controller tried to speak too slow, figure 23 describes the probable relationship of voice speed to recognition accuracy and acceptance. This interpretation of recognizer performance shows the exponential decay of recognition accuracy with slow speech due to over-articulation and the exponential decay of recognition acceptance with fast speech due to the slurring of words.

While the unit missed many long phrases, the phrases it did accept were recognized with little error. Also, even though the controller carried on conversations with others during the tests, the unit never accepted any input other than aircraft calls. This is quite an event to witness. At times, the controller would talk for five

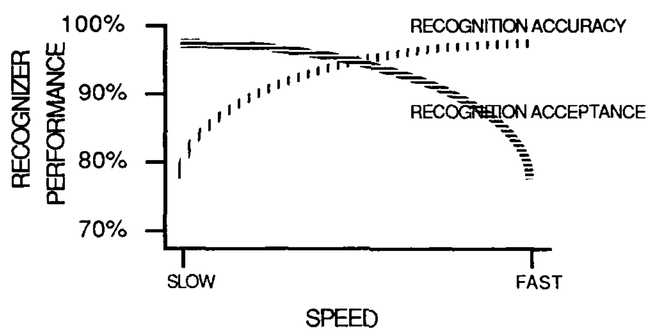

Figure 23. Voice speed versus recognition accuracy and acceptance.

minutes to someone else with the unit waiting for correct voice input. When he had to interrupt his conversation to answer the radio call of an aircraft, the unit would respond with the correct parsing and begin waiting for another correctly voiced phrase.

With knowledge of the operational performance of this prototype voice recognizer system, it is possible to develop applications to enhance the effectiveness and performance of air traffic controllers. This is the subject of our next chapter.

Chapter Five

When the parsing capabilities of the voice recognizer are combined with the ability to customize the output of the recognizer, a wide variety of applications can be developed. This chapter explores possible applications of the voice recognizer in an Air Traffic Control environment. The design and suitability of each application will be discussed.

APPLICATIONS

As mentioned in Chapter Four, the Voice Planner™ software processes the grammar definition file into a binary file which the voice recognizer uses to determine what voice patterns are needed from the user. This binary file is also used to define the sequence in which the recognizer will parse a sentence. Once a sentence is parsed, the recognizer sends an ASCII character string through a RS-232 terminal connection to the host computer. The form of the ASCII character string is determined by another text file called a translation table file.

The translation table file determines what ASCII characters will be sent for each word match in a sentence. By default, the recognizer will send the ASCII equivalent of each word with a single space separator. Also, the default option has no initiator string and a carriage return as the terminator for each parsed phrase. Thus, the default option sends an ASCII string which looks as if someone had typed the sentence from a keyboard.

In many applications, the user will want to change the recognizer output so it will conform to the input requirement of some postprocessing software, to create special functions for particular words, or to just present a more readable output. To make the recognizer output more readable during tests, the following translation table file was used in the development of the ground controller recognition grammar:

```
||              ; no initiator string
|040            ; space separator
|015            ; carriage return terminator

zero            0
one             1
two             2
three           3
four            4
five            5
six             6
seven           7
eight           8
nine            9
```

GROUND CONTROLLER TRANSLATION TABLE FILE

This translation file changed all digits from a spelled-out to a numeric form.

The general architecture for an ATC tower application is shown in figure 24. Each controller position has a VocaLink® voice recognizer. The three recognizers can be attached to dedicated computers or networked to a host computer running postprocessing software. The postprocessing software can include applications for switching, voice data storage, and expert systems. Before the description of each type of application is presented, the method used to communicate the output of the voice recognizer to the Standard 286 AT compatible computer will be detailed.

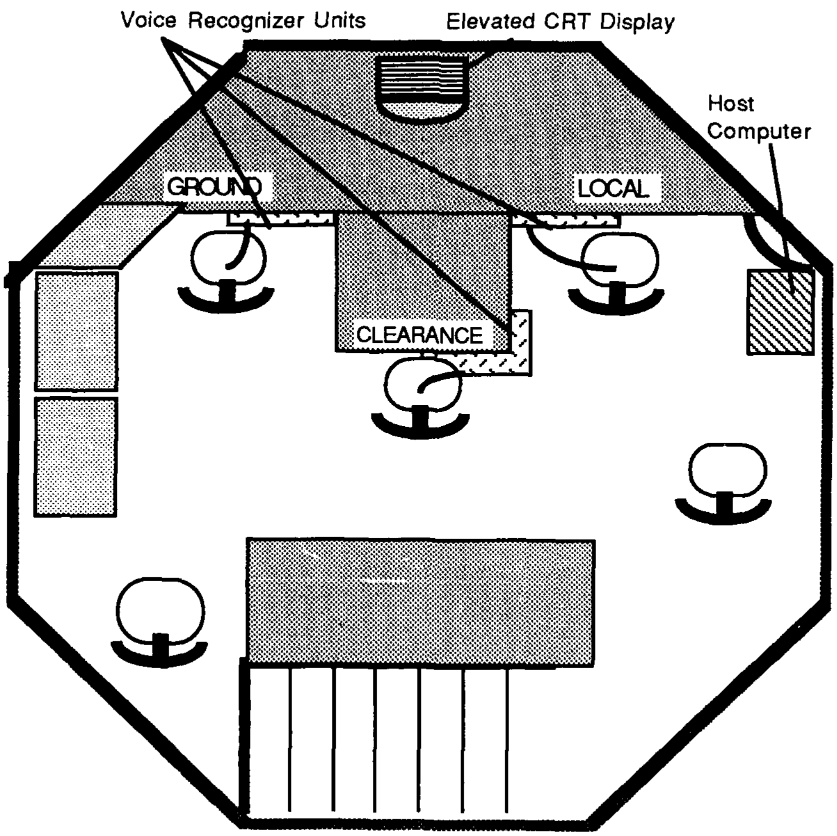

Figure 24. ATC Application Architecture.

The recognizer output is communicated to the Standard 286 through a serial port. The COM1 serial port in the 286 was used to receive input. The Standard 286 ran a Turbo PROLOG postprocessing program to receive and display recognizer output. This program appears in Appendix D. The program allowed recognizer information to be communicated to a PROLOG environment. Thus, the following applications specify Turbo PROLOG functions which will be used to accomplish particular functions.

A Switching Device

A switching device application uses the convenience of voice to accomplish manual tasks. In this application, the voice recognizer would send a series of ASCII codes to the host computer (Standard 286, in this case) in order to call programs to run emergency and procedural checklists.

In the modified ground controller recognition grammar, the phrase "runway-change" was included to determine the reliability of one word commands. It should be noted that the recognizer treats the phrase "runway-change" as one word, i.e. one voice pattern. As discussed in Chapter Four, the reliability and acceptability of select one word commands is very high. Similar results with other select words would be expected. Notice that the particular command must be selected and tested beforehand, because some words may be unusually difficult to parse. For instance, poor recognition acceptance with the single word "roger" was noted.

The translation table file would modify the ASCII output of the phrase "runway-change" to an acceptable PROLOG environment command such as:

sample translation file entry for DOS system calls:	
runway-change	system(b r.txt)

This system command would display the text file 'r.txt' consisting of a runway-change checklist as in figure 25 on a conveniently placed monitor.

PROPOSED RUNWAY CHANGE	
30 Minutes Prior	SOFs notify appropriate RSUs and PHX of proposed runway change time. T-38 SOF will notify ABQ and tower with change time. Tower will coordinate with Fire Dept and Base Ops to insure barriers get changed.
20 Minutes Prior	Hold all aircraft in the chocks. Discontinue radar departures.
10 Minutes Prior	Discontinue all take-offs from and straight-ins to the center runway. Change center runway barriers. (Dual runway ops) Back-taxi any A/C down the center runway. Gasser will confirm runway change with Mojack

Figure 25. Screen display for checklist.

Voiced Data Collection

Once a computer has captured voiced data, it can process that data in a number of ways. It could, for instance, store it or transmit it to another location. In the application of storage, the voice data could be collected into a PROLOG dynamic database. The translation table file to accomplish this might look like:

<u>sample translation file entry for PROLOG database:</u>

```
assertz( obj(        ; initiator string
,                    ; separator
) ).                 ; terminator

awol                 "awol"
willy                "willy"
one                  "1"
two                  "2"
runway               "runway"
taxi                 "taxi"
```

Thus, the recognizer output for the phrase "awol 21, taxi runway 12" would be: *assertz(obj("awol","2","1","taxi","runway","1","2")).* The assertz predicate allows information to be stored in a PROLOG database as true. Users running the program would have the full power of the PROLOG language to draw information from this database.

In the application of transmission, the voiced data could be transmitted to other ground agencies. This would be especially useful in emergency situations. In the area of international operations, the parsed language of a foreign controller could be transmitted directly to the aircraft. This would be helpful in many countries where air traffic controllers are hard to understand and could possibly alleviate the requirement for foreign controllers to learn English.

The suitability of voice recognition in an ATC data storage or transmission application is suspect given the current recognition acceptance rate for long phrases. Recognition acceptance would have to be in the high 90's for these applications to be useful. A controller could not be expected to repeat a phrase in an operational environment unless it occurred infrequently.

Expert Systems

The ability to process voiced information can be expanded into the area of expert systems. With the ability to collect data into a PROLOG dynamic database, you may also compare this data to a set of procedure rules. The procedure rules would take the form of clauses in the Turbo PROLOG language. An example of what these clauses would look like for a local controller application could be:

```
predicates
        warning
        takeoff(callsign,time,status,runway,wind,altimeter)
        landing(callsign,time,status,runway,wind,altimeter)

clauses
        warning if
                takeoff(_,_,cleared,21,_,_) and
                landing(_,_,cleared,21,_,_).
```

The clause "warning" would become true if an aircraft were to be cleared onto runway 21 while another is cleared for takeoff. When the "warning" clause becomes true, a tone could alert the controller to a possible aircraft conflict.

As with the voiced data application, the suitability of an expert system application becomes suspect when considering the recognition acceptance of long phrases. Obviously, unless the recognition acceptance and reliability is very high, the data which the procedure rules compare becomes invalid. With invalid data, the expert system is not useful.

Chapter Six

CONCLUSION

Within the Air Traffic Control environment the application of voice recognition is limited by the difficulty in parsing long phrase sentences. This difficulty is not in the recognition reliability once the voice recognizer accepts input, but in the requirement to repeat many long phrases in order to have the unit accept the phrase. The prototype equipment gave extremely high recognition accuracy and rarely produced a false recognition.

Voice recognition applications which can be expected to produce reliable systems are manual switching, data collection, and expert systems development with the use of single and short (up to seven words) phrases. The parsing of long phrases is currently limited by recognition acceptance. It should be noted that the prototype unit was using 8086 microprocessor technology and with the advent of 286 and 386 microprocessor designs, recognition acceptance should improve.

Under current voice recognition technology, there seems to be a tradeoff between high accuracy and high acceptance. For instance, some voice recognition systems currently on the market produce a high number of false recognitions. This is due to their voice pattern matching algorithm which may require that some voice pattern match be made no matter how remote. To guard against false recognitions and ensure high accuracy, the tolerance for accepting a particular voice pattern must be set high. This results in a greater number of word rejections and decreases recognition acceptance.

It should be noted that these results apply to an ATC environment and that friendlier environments should produce better results. For instance, an environment which allowed users to repeat a phrase if not accepted by the recognition unit and which trained users to speak at a moderate pace could expect greater recognition acceptance of long phrases. In any event, this environment would allow the user to correct accuracy and acceptance errors.

The presentation of the sentence diagramming technique should assist others in the establishment of complex grammars for voice recognizers. As technology increases in the voice recognition area, requirements to establish more and more complex grammars will appear. The sentence diagramming technique can be used to map the structure of the grammars which will run on these advanced systems. While borrowing from verb-centered semantic approaches, the technique is flexible enough to allow application to a wide variety of areas. By including a word sequencing scheme, the sentence diagramming technique can be directly coded into a grammar definition for voice recognizers.

FURTHER RESEARCH

Voice recognizers are dependent on their grammar definition to parse voice data. Therefore, further effort in optimizing the grammar might yield better recognition acceptance. One way to reduce grammar complexity would be to reduce the number of optional phrases available to a user. Schemes might also be developed where only the use of short phrases is needed in a particular application or the repeating of a long phrase is not as detrimental to the user as in a control tower application.

Within the area of recognizer performance, an investigation into the speed of a user's voice and its effect on recognition accuracy and recognition acceptance is warranted. If one were to attempt to quantify the point at which maximum accuracy and acceptance occurs for the Vocalink® S4000 in an ATC environment, it would appear to occur at a point between two and three words per second. A thorough study could establish this voice speed for the Vocalink® S4000 and similiar recognition units.

In addition, this voice speed limit seems dependent not on the processing power of the unit, but on the quality of the utterance by the human. In other words, a speed is reached at which word sounds must be combined or slurred in order maintain the rate of speech. The increased rate of speech is usually not difficult for a human, who translates the speech information phonetically, to understand. However, it could be very difficult for a machine, who translates the speech information acoustically, to match.

The use of a voice recognizer as a switching device is somewhat limited in a control tower, but could be more practical in a ground radar environment. In this environment, a controller must operate machinery and talk to aircraft at the same time. A voice recognizer with the appropriate machine interface could assist the user with machine control. The machine control application could include the expanding or condensing of the radar area portrayed on the radar scope, the control of scope brightness and constrast, or the input of aircraft data for a particular target.

The application of voice recognition in the ground radar environment could also provide for limited data collection as in the input of aircraft information into a central computer. Applications could also include expert systems as in the issuance of a warning when the controller clears an aircraft below a minimum safe altitude.

References

Bailey, R. W. (1982). *Human performance engineering: A guide for systems designers.* Englewood Cliffs: Prentice-Hall, Inc.

Burnham, W. D. and Hall, A. R. (1985). *Prolog programming and applications.* New York: John Wiley & Sons.

Cooper, M. (1987). Human factor aspects of voice input/output. *Speech Technology,* 3(3), 82-86.

Fillmore, C. J. (1968). The case for case. *Universals in linguistic theory.* (pp. 1 - 89). New York: Holt, Rinehart and Winstor, Inc.

Kantowitz, B. H. and Sorkin, R. D. (1983). *Human factors: Understanding people-system relationships.* New York: John Wiley & Sons.

Kunz, J. R. (1982). *The American Medical Association family medical guide.* New York: Random House.

Lea, W. A. (1979). *Trends in speech recognition.* New York: John Wiley & Sons.

Parsons, T. (1987). *Voice and speech processing.* New York: McGraw-Hill Book Company.

Rosch, W. L. (1987). Voice recognition: Understanding the master's voice. *PC Magazine,* 6(18), 261-308.

Simpson, C. A. and Ruth, J. C. (1987). The phonetic discrimination test for speech recognizers: Part I. *Speech Technology,* 3(3), 48-53.

Simpson, C. A. and Ruth, J. C. (1987). The phonetic discrimination test for speech recognizers: Part II. *Speech Technology,* 3(4), 58-61.

Somers, Harold L. (1987). *Valency and case in computational linguistics.* Edinburgh: University Press.

Sowa, John F. (1984). *Conceptual structures: Information processing in mind and machine.* Reading: Addison-Wesley Publishing Co.

Teja, E. R. and Gonnella, G. (1983). *Voice technology.* Reston: Reston Publishing Company, Inc.

Thomas, T. J. (1987). Speech recognizer assessment in the laboratory, not in the field. *Speech Technology,* 3(3), 89-93.

White R. W. and Parks D. L. (1985). *Study to determine potential flight applications and human factors design guidelines for voice recognition and synthesis systems.* Boeing Commercial Airplane Company: Langley Research Center, National Aeronautics and Space Administration, Contractor Report 172590.

Appendix A - Design Guidelines/Specifications

GENERAL-PURPOSE DESIGN GUIDELINES/SPECIFICATIONS

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

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

- Direction and interrogation messages associated with each task must be defined.

- Message syntax should be noted and variable syntax combinations considered.

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

- Necessary response time should be appraised.

- With recognition system controller --> first- and second-choise words and associated ratings must be available for the recognition system.

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

- Interface hardware requirements must be defined.

- Interface software requirements must be defined.

- Environmental conditions must be defined.

- Storage requirements for voice records must be considered and planned.

Appendix B - Controller Transcripts

LOCAL, GROUND, AND CLEARANCE DELIVERY TRANSCRIPTS

hr/min/sec Local Controller Transcript (29 Jan '88 : 0930-1230)

Call sign	Aircraft type
AWOL	T-38
Hook	T-37
Hungry	F-5
Spydr	F-5

0000	Aircraft holding short of 30 center say your call sign again.
0005	AWOL 44, Willy tower, wind calm, cleared for takeoff, change to departure.
0026	AWOL 35, tower, hold short.
0203	AWOL 35, taxi across runway 30 center.
0443	AWOL 46, Willy tower, wind calm, cleared for takeoff, change to departure.
0458	Willy 80, Willy tower, hold short.
0620	AWOL 42, tower, report final approach fix, 30 center, wind calm, altimeter 3025.
0633	Willy 80, wind calm, cleared for takeoff, change to departure, traffic T38 at 7 miles.
0713	AWOL 42, wind calm, cleared for touch-and-go runway 30 center, traffic T38 departure position.
0831	AWOL 44, Willy tower, report final approach fix runway 30 center, wind calm, altimeter 3025.
0953	AWOL 44, wind calm, cleared for the option runway 30 center, your traffic in a Arie departure midfield center runway.
1011	Willy 51, Willy tower, hold short traffic on final.
1022	51 roger.
1102	Willy 51, wind calm, cleared for takeoff, change to departure.
2637	AWOL 42, Willy tower, report final approach fix runway 30 center, wind calm, altimeter 3025.
2753	AWOL 42, wind calm, cleared tough-and-go runway 30 center.
2823	AWOL 43, Willy tower, report final approach fix 30 center, wind calm, altimeter 3025.
2945	AWOL 56, Willy tower, hold short, traffic on final.
2958	AWOL 43, check wheels down, wind calm, cleared for the option runway 30 center.
3127	AWOL 56, wind calm, cleared for takeoff, change to departure.
3404	AWOL 36, Willy tower, report final approach fix runway 30 center, wind calm, altimeter 3025.
3503	AWOL at 10 miles, report final approach fix 30 center, wind calm, altimeter 3025, your number 2 to follow a T-38 at 5-and-a-half miles.
3518	AWOL 64, tower, taxi across 30 center without delay.
3527	AWOL 36, wind calm, cleared for the option runway 30 center.

3620 AWOL 27, wind calm, cleared to land 30 center.
3940 AWOL 54, Willy tower, wind calm, cleared for takeoff, change to departure.
4156 T37 calling tower, repeat call sign, Hook 35?
4202 Hook 35, Willy tower, wind calm, cleared for takeoff, change to departure.
4323 AWOL 36, Willy tower, report final approach fix runway 30 center, wind calm, altimeter 3025.
4509 AWOL 36, wind calm, cleared to land runway 30 center.
4533 AWOL 66, Willy tower, hold short, traffic inbound.
4643 61, roger, hold short.
4645 AWOL 66, Willy tower, wind calm, cleared for takeoff, change to departure.
4745 AWOL 36, contact SOF when off runway.
4844 AWOL 61, wind calm, cleared for takeoff, change to departure.
5010 Hungry 1, roger, hold short of runway, you're coming in real garbled.
5210 Hook 35, Willy tower, report final approach 30 center, wind calm, altimeter 3025.
5335 Hungry 1, Willy tower, hold short traffic on final.
5421 Hook 35, wind calm, cleared for the option runway 30 center, unable kilo for a F5 going off the center runway.
5545 AWOL 75, tower, hold short.
5627 AWOL 75, taxi across runway 30 center.
5651 Hook 35, frequency change approved.
5653 Hungry 1, taxi into position and hold runway 30 center, awaiting an ATC release from Pheonix.
5806 Hungry 1, say again.
5810 We're waiting for a call back from Pheonix, sir.
5814 Hungry 1, roger, wind calm, cleared for takeoff.
5850 AWOL 63, tower, hold short.
5917 AWOL 63, taxi into position and hold 30 center.
5956 AWOL 63, wind calm, cleared for takeoff, change to departure.

1:0117 AWOL 70, tower, change to departure, wind calm, cleared for takeoff 30 center.
1:0404 Hook 35, Willy tower, report final approach fix runway 30 center, wind calm, altimeter 3025.
1:0445 Spydr 21, roger, are you ready for departure.
1:0449 Spydr 21, roger, taxi into position and hold, separate from air departure off the outside.
1:0524 Spydr 21, wind calm, cleared for takeoff 30 center, traffic T-37 approaching final approach fix, frequency change approved your decretion.
1:0544 Hook 35, roger, wind calm, cleared for the option 30 center, traffic flight of two F-5s departure position, caution wake turbulence.
1:0618 Willy 80, tower, report final approach fix 30 center, wind calm, altimeter 3025, number two following a T-37 approximately 4 miles.
1:0631 35, check wheels down, cleared for the option.
1:0724 Roger, Hook 35, frequency change approved.
1:0807 Willy 80, wind calm, cleared for low approach.
1:0820 Willy 51, report final approach fix 30 center, wind calm, altimeter 3025.

1:0904	Egret 1, Willy tower, hold short.
1:0943	51, wind calm, cleared for low approach, contact gas 2 mile final.
1:0955	Hook 44, Willy tower, hold short.
1:1047	AWOL 73. Willy tower, hold short.
1:1056	AWOL 11, roger, hold short.
1:1102	Egret 1, change to departure, wind calm, cleared for takeoff.
1:1223	Hook 44, change to departure, wind calm, cleared for takeoff, caution wake turbulence.
1:1356	AWOL 73, Willy tower, taxi into position and hold, awaiting separation from T-37 departure.
1:1410	AWOL 64, Willy tower, report 5 miles 30 center, wind calm, altimeter 3025.
1:1446	AWOL 73, change to departure, wind calm, cleared for departure 30 center.
1:1537	AWOL 64, wind calm, cleared for touch-and-go 30 center.
1:1620	Willy 80, Willy tower, report final approach fix 30 center, wind calm, altimeter 3025.
1:1714	AWOL 11, change to departure, wind calm, cleared for takeoff, traffic T-38, full stop at Kipler.
1:1751	Willy 80, wind calm, cleared to land runway 30 center.
1:2222	Hook 44, Willy tower, report final approach fix 30 center, wind calm, altimeter 3025.
1:2359	Hook 44, wind calm, cleared for the option, 30 center.
1:2425	Post 12, hold short traffic on final.
1:2643	Hook 44, frequency change approved.
1:2710	Post 12, roger, standby I'm waiting for separation.
1:2718	Post 12, taxi into position and hold.
1:2742	Post 12, change to departure, wind calm, cleared for takeoff.
1:3110	Hook 23, Willy tower, report final approach fix 30 center, wind calm, altimeter 3025.
1:3130	Pipeline 057, Willy tower.
1:3141	Pipeline 057, roger, enter the Williams ATA from the west at or below 2000 feet. Runway 30 in use, Williams altimeter 3025, wind is calm.
1:3200	Pipeline 057, understand you're west now.
1:3207	Roger.
1:3230	Hook 23, wind calm, cleared to land, 30 center.
1:3243	AWOL 15, Willy tower, hold short, traffic on final.
1:3412	AWOL 17, roger, hold short.
1:3452	Post 16, roger.
1:3455	AWOL 15, change to departure, wind calm, cleared for takeoff.
1:3601	AWOL 75, Willy towr, report final approach fix 30 center, wind calm, altimeter 3025.
1:3624	AWOL 17, change to departure, wind calm, cleared for takeoff 30 center.
1:3630	Pipeline 057, I observe you exiting the Williams ATA, frequency change approved.
1:3640	Willy 81, Willy tower, hold short.
1:3651	Page 66, Willy tower report final approach fix 30 center, wind calm, altimeter 3025.
1:3718	AWOL 75, wind calm, cleared for the option 30 center.

1:3807	Page 66, are you at post now.
1:3815	Roger.
1:3842	Post 16, change to departure, wind calm, cleared for takeoff, traffic T-37 final approach fix.
1:3902	Page 66, wind calm, cleared the option 30 center, traffic T-38 departure position.
1:3959	Willy 81, wind calm, cleared for takeoff, traffic T-37 3 miles.
1:4153	Page 66, frequency change approved.
1:4251	AWOL 15, Willy tower, report 5 miles 30 center, wind calm, altimeter 3025.
1:4352	AWOL 17, Willy tower, report 5 miles 30 center, wind calm, altimeter 3025, number 2 to follow same type, final approach fix.
1:4406	AWOL 15, check wheels down, wind calm, clear for the option.
1:4515	AWOL 17, wind calm, cleared for touch and go, 30 center.
1:5104	Post 22, Willy tower, change to departure, wind calm, cleared for takeoff 30 center.
1:5215	Spydr 24 Willy tower, proceed across the center runway, ????.
1:5235	Spydr 24, roger.
1:5250	Willy 1, Willy tower, report 5 miles 30 center, wind calm, altimeter 3025.
1:5408	Willy 1, wind calm, cleared for the option 30 center.
2:0308	AWOL 26, Willy tower, change to departure, cleared for takeoff.
2:0324	Hook 36, Willy tower, report final approach 30 center, wind calm, altimeter 3025.
2:0443	AWOL 31, change to departure, wind calm, cleared for takeoff, traffic a T-37 final approach fix.
2:0510	Hook 36, wind calm, cleared for option 30 center.
2:0757	Hook 47, Willy tower, report final approach fix 30 center, wind calm, altimeter 3021.
2:0810	36, frequency change approved.
2:0935	30, Willy tower, change to departure, wind calm, cleared for takeoff, traffic T-37 5 mile final.
2:0949	47, wind calm, cleard for touch and go.
2:1204	AWOL 26, Willy tower, reprot 5 mile final runway 30 center, wind calm, altimeter 3021.
2:1213	Hook 47, frequency change approved.
2:1317	AWOL 25, wind calm, cleared for touch and go.
2:1332	AWOL 31, Willy tower, report 5 mile final runway 30 center, wind calm, altimeter 3021.
2:1450	AWOL 31, wind calm, cleared for touch and go.
2:1539	Calling 10 mile final, report final approach fix 30 center, wind calm, altimeter 3021.
2:1545	50 roger.
2:1608	Post 34, Willy tower, change to departure, wind calm, cleared for takeoff.
2:1739	50, wind calm, cleared for touch and go.
2:1812	AWOL 15, Willy tower, report 5 mile final 30 center, wind calm, altimeter 3021.
2:1940	AWOL 15, wind calm, cleared for touch and go.
2:1948	50, frequency change approved.

2:1955	Hook 71, hold short, traffic on final.
2:2004	AWOL 17, Willy tower, report 5 mile final 30 center, wind calm, altimeter 3021.
2:2123	AWOL 17, wind calm, cleared for touch and go.
2:2130	Roger.
2:2247	Hook 71, change to departure, wind calm, cleared for takeoff.
2:2330	Post 32, Willy tower, hold short.
2:2341	66, hold short.
2:2430	AWOL 33, hold short.
2:2432	Hook 66, change to departure, wind calm, cleared for takeoff.
2:2522	Hook 55, report final approach fix 30 center, wind calm, altimeter 3021.
2:2636	Hook 32, change to departure, wind calm, cleared for takeoff.
2:2722	Hook 55, wind calm, cleared for touch and go.
2:2801	Hook 40, Willy tower, hold short.
2:3000	55, frequency change approved.
2:3005	AWOL 11, report 5 mile final 30 center, wind calm, altimeter 3021.
2:3013	AWOL 33, change to departure, wind calm, cleared for takeoff.
2:3124	AWOL 11, wind calm, cleared to land.
2:3251	Hook 71, Willy tower, report final approach fix 30 cneter, wind calm, altimeter 3021.
2:3300	Hook 40, change to departure, cleared for takeoff.
2:3445	71, wind calm, cleared for touch and go.
2:3608	71, check wheels down.
2:3713	Hook 71, frequency change approved.
2:3854	AWOL 33, Willy tower, report 5 mile final 30 center, winds 060 at 6, altimeter 3021.
2:3951	AWOL 33, winds 050 at 6, cleared for touch and go.
2:4303	Willy 81, Willy tower, report 5 mile final runway 30 cneter, winds 030 at 4, altimeter 3021.
2:4429	Willy 81, winds 100 at 4, cleared to land.
2:4553	Hook 14, hold short.
2:4608	Hook 14, change to departure, winds 090 at 4, cleared for takeoff.
2:4717	Page 97, Willy tower, report final approach fix to runway 30 cneter, wind 060 at 4, altimeter 3021.
2:4928	AWOL 26, Willy tower, report final approach fix runway 30 center, winds 080 at 6, altimeter 3021, number 2 to follow a T-37 four-and-one-half mile final.
2:4948	Page 97, wind 080 at 6, cleared for low approach.
2:5000	Page 97, copy.
2:5010	AWOL 46, Willy tower, shold short of runway, traffic on final.
2:5050	AWOL 26, winds 080 at 4, cleared for touch and go.
2:5116	Page 97, frequency change approved.
2:5222	AWOL 46, change to departure, wind 070 at 8, cleared for takeoff.
2:5250	AWOL 51, Willy tower, hold short of runway.
2:5315	AWOL 43, Willy tower, roger.
2:5320	43, roger.
2:5417	AWOL 51, change to departure, wind calm, cleared for takeoff.
2:5422	Roger, cancel takeoff clearance, continue holding short.

2:5437	43, change to departure, wind 100 at 4, cleared for takeoff.
2:5629	AWOL 51, change to departure, wind 070 at 7, cleared for takeoff.
2:5641	Hook 13, hold short.
2:5757	Hook 13, change to departure, wind 090 at 6, cleared for takeoff.
=====>	3:0000

hr/min/sec Ground Controller Transcript (29 Jan '88 : 1245-1345)

Call sign	Aircraft type
AWOL	T-38
Hook	T-37
Hungry	F-5
Spydr	F-5

0042	AWOL 51, Willy ground, taxi 30 center, wind calm, altimeter 3023.
0238	296, advice termination.
0440	AWOL 55, continue taxi and hold short of the center runway, tower 255.6 when ready.
0827	AWOL 46, Willy ground, taxi 30 center, wind calm, altimeter 3023.
0939	AWOL 50, Willy ground, taxi runway 30 center, wind calm, altimeter 3023.
1139	AWOL 56, Willy ground, taxi 30 center, wind calm, altimeter 3023.
1157	474, Willy ground, advise termination.
1633	AWOL 54, Willy ground, taxi 30 center, wind calm, altimeter 3023.
1656	Hook 24, ground, taxi 30 left, wind calm, altimeter 3023.
1827	Hook 26, taxi 30 left, wind calm, altimeter 3023.
1850	Hook 33, taxi 30 left, wind calm, altimeter 3023.
2323	Roger.
2418	Hook 32, taxi 30, wind calm, altimeter 3023.
2641	Roger, taxi approved.
2648	Roger, wind calm, altimeter 3023.
2708	25, taxi approved.
2715	Roger, taxi approved.
2719	T38 flying status, supervise solo.
2741	61, taxi 30, wind calm, altimeter 3023.
2926	Roger, taxi 30, wind calm, altimeter 3023.
2939	Roger.
3042	With the numbers, taxi approved.
3504	63, taxi 30, wind calm, altimeter 3023.
3532	Roger, taxi 30, wind calm, altimeter 3023.
3743	Willy altimeter 3024.
3746	Taxi 30, wind calm, altimeter 3024.
3944	Roger, advise termination, outside air temperature is plus 51.
4056	Taxi 30, wind calm, altimeter 3024.
4736	Reno 06, taxi 30, wind calm, altimeter 3024.
5230	Hook 51, taxi 30, wind calm, altimeter 3024,
5403	Hook 46, taxi 30, wind calm, altimeter 3024.
5503	72, ground, taxi approved.
5516	74 flight, taxi approved.
5633	Taxi 30, wind calm, altimeter 3024.
5647	AWOL 73, ground, taxi approved, wind 020 at 6.
5802	New Williams altimeter, 3023.
5830	Willy 91, taxi approved.
5834	Willy 91, do you got your clearance.
5858	Calling for taxi, taxi 30, wind calm, altimeter 3023.

5906	91, ground, say your proposed departure time.
5935	866, ground.
5939	Roger, clearance delivery 289.4, frequency change approved, report back.
====>	1:0000

hr/min/sec Ground Controller Transcript (19 Feb '88 : 1400-1600)

Call sign	Aircraft type
AWOL	T-38
Hook	T-37
Hungry	F-5
Spydr	F-5

0002	Calling for taxi, taxi 30 center, wind 250 at 6, altimeter 2984, say your call sign again.
0010	Hook 27.
0047	Ground radio check, your loud and clear.
0158	Reno 7, ground.
0202	Reno 7, squawk 4211.
0207	Reno 7, correction, your squawk is 4241.
0410	38223, roger, good day.
0420	8230, your loud and clear.
0515	Willy 91, Willy ground, taxi runway 30 center, wind calm, altimeter 2984.
0814	Hook 36, Willy ground, taxi runway 30 left, wind calm, altimeter 2984.
0825	29("niner")84.
0946	Hook 30, Willy ground, taxi runway 30 left, wind 190 at 5, altimeter 2984.
1106	Hook 35, Willy ground, taxi approved.
1113	AWOL 63, Willy ground, taxi runway 30, wind calm, altimeter 2984.
1125	Hook 34, roger.
1220	AWOL 61, Willy ground, taxi approved.
1306	4379, roger, report termination.
1409	42, Willy ground, taxi runway 30, wind calm, altimeter 2984.
1455	T37 flying status is dual.
1725	Hook 24, Willy ground, taxi runway 30 left, wind calm, altimeter 2984.
1834	4379, roger, good day.
1842	AWOL 54, Willy ground, taxi approved.
2042	Hook 24, frequency change approved.
2150	Hook 37, roger.
2331	Hook 35, Willy ground, taxi runway 30 left, wind calm, altimeter 2983.
2550	AWOL 54, frequency change approved, report back.
2647	AWOL 65, roger.
2732	AWOL 71, Willy ground, taxi runway 30 center, wind calm, altimeter 2983.
3051	AWOL 13, Willy ground, taxi approved.
3123	Hook 33, say again.
3128	33, Willy ground, taxi runway 30, wind calm, altimeter 2983.

3313 41, Willy ground, taxi runway 30 left, wind calm, altimeter 2983.
3510 Calling for taxi, taxi runway 30 left, wind calm, altimeter 2983.
3517 Hook 40.
3636 Willy 57, Willy ground, taxi runway 30 center, wind calm, altimeter 2983.
3923 Willy 81, Willy ground, taxi to runway 30 center, wind calm, altimeter 2983.
4049 AWOL 77, Willy ground, taxi approved.
4657 Hook 33, Willy ground, taxi runway 30 left, wind 240 at 6, altimeter 2983, bird condition yellow.
4853 AWOL 71, Willy ground, roger, taxi approved.
5018 Willy 55, taxi runway 30 center, wind 270 at 6, altimeter 2983, bird condition yellow.
5238 Roger, say position.
5242 Roger, frequency change approved, report back up.
5247 AWOL 15, taxi approved, bird condition yellow.
5633 Altimeter now 2982.
5652 Altimeter 2982.
5824 Hook 51, taxi runway 30 left, wind calm, altimeter 2982, bird condition yellow.
1:0323 Hook 72, taxi runway 30 left, wind calm, altimeter 2982, bird condition yellow.
1:0336 Taxi 30 center.
1:0348 Roger, advice termination.
1:0537 AWOL 22, Willy ground, taxi runway 30 center, wind calm, altimeter 2982, bird condition yellow.
1:0601 Loud and clear.
1:1015 Hook 54, Willy ground, taxi runway 30 left, wind calm, altimeter 2982, bird condition yellow.
1:1029 AWOL 16,Willy ground, taxi approved.
1:1058 Roger.
1:1114 Roger, taxi approved.
1:1129 60, taxi approved.
1:1144 AWOL 11, taxi approved.
1:1152 Chandler PIREP, a T-37 at flight level 230, 19 miles to the west of Chandler, base of the clouds are unknown with overcast tops at flight level 190, icing and turbulence were negative, time of report 2155 zulu. At 2203 zulu, a T-37, 3 miles south-south-east of Chandler at 15 thousand, there's light rime icing.
1:1232 804, Willy ground, taxi runway 30 left, wind calm altimeter 2982.
1:1310 Advise termination.
1:2017 Tall 95, my ????? shows spot 93 your left departure end.
1:2045 Tall 95, proceed across runway 30 left, departure clear.
1:2103 And Tall 95, are you familiar with transient parking.
1:2151 Advice termination.
1:2158 Roger.
1:2251 11, tower.
1:2300 We weren't reported anything.
1:2303 Is he a T-37.

1:2310	Willy 11, ground.
1:2314	Where is he parked at.
1:2326	Roger, he's taxing back.
1:2912	Hook 64, Willy ground, taxi runway 30 left, winds 270 at 4, altimeter 2982, bird condition yellow.
1:2935	Hook 64, roger.
1:3241	Roger.
1:3307	428, Willy ground, hold short of runway 30 left, ??? center.
1:3537	428, proceed across runway 30 left, ??? center, make a half right turn while crossing to the diagonal taxiway.
1:3602	428, the follow-me truck ahead of you will assist you in parking.
1:3713	Hook 72, you're loud and clear.
1:3745	Advise termination.
1:3840	Hook 63, Willy ground, taxi runway 30 left, wind 300 at 8, altimeter 2982, bird condition yellow.
1:3852	Hook 63, did you copy.
1:3908	672, Willy ground.
1:3915	672-2, ground.
1:3921	672-2, Willy ground.
1:3935	672-2, Willy ground.
1:3941	672, I don't think he can hear me.
1:3948	Hook 65, taxi runway 30 left, winds 300 at 6, altimeter 2982, bird condition yellow.
1:4011	Roger.
1:4048	Chandler PIREP, at 2203 zulu, a T-37, 3 miles to the south-south-east of Chandler at 15 thousand reported light rime icing. Another T-37 at flight level 230, 19 miles to the west of Williams reported base of clouds unknown with overcast tops at flght level 190, icing and turbulence were negative, that was reported at 2155 zulu.
1:4443	Hook 26, taxi 30, wind 330 at 4, altimeter 2982.
1:4809	Hook 72, taxi 30, wind calm, altimeter 2982.
1:4911	Hook 73, taxi 30 center, wind calm, altimeter 2982.
1:5224	Back-taxi approved.
1:5434	Calling for taxi, taxi 30, wind 300 at 6, altimeter 2982.
1:5455	Post 36, taxi approved.
1:5907	Hook 74, taxi 30, wind calm, altimeter 2981.
=====>	2:0000

hr/min/sec Clearance Delivery Transcript (29 Jan '88 : 1400-1440)

Call sign	Aircraft type
AWOL	T-38
Hook	T-37
Hungry	F-5
Spydr	F-5

0312	Page 66, Willy clearance delivery, you're cleared to the Fort Huachuka airport, Kilo departure, maintain 15 thousand, departure frequency will be 317.5, squawk 4215.
0757	Willy 91, Willy clearance delivery, you're cleared Williams airport, shark 7, maintain flight level 190, departure frequency will be 3175, squack 4230.
1218	Reno 4, Willy clearance delivery, you're cleared to Williams airport, Reno 2, maintain flight level 270, departure frequency 317.5, squawk 4217.
1239	Willy 07, Willy clearance delivery, you're cleared to Norton airport, Stellar 5 departure then as filed, maintain flight level 190, expect flight level 370, sqauwk 0701.
1301	Willy 49, standby, you're clearance is on request.
1346	Willy 49, clearance.
1348	Willy 49, base ops is submitting you're flight plan right now, the delay is unknown.
2000	Willy 49, Willy clearance delivery, you're cleared to Port Mague airport, Peter 6 departure, Drake transition direct Needles then as filed, maintain flight level 190, expect flight level 350, squawk 2652.
2523	Is your call sign Page 84.
2525	Page 84, Willy clearance delivery, you're cleared to Fort Huachuka airport, Squeak 1, maintain 12 thousand, expect 15 thousand 10 minutes after departure, departure frequency will be 363.0, squawk 4371.
2755	Page 18, Willy clearance delivery, you're cleared to Fort Huachuka airport, Squeak 1 departure then as filed, maintain 12 thousand, expect15 thousand 10 minutes after departure, departure frequency will be 363.0, squawk 4370.
2831	Page 18, you're initial altitude to maintain is 12 thousand.
====>	4000

hr/min/sec Clearance Delivery Transcript (19 Feb '88 : 1100-1320)

Call sign	Aircraft type
AWOL	T-38
Hook	T-37
Hungry	F-5
Spydr	F-5

0942	Syder 21, your clearance is on request, expect lengthy delay, approximately 30 minutes.
1000	Syder 21, roger.
1018	07, Willy clearance, cleared to Jody, superstition niner departure, then as filed, maintain flight level 270, departure frequency will be 317.5, squawk 0716.
1034	07, you copied cleared to Jody?
3238	Reno 04, Willy clearance delivery, standby, clearance on request, expect lengthy delay, the computer is down, advice T-37 or T- 38.
3255	T-38, roger.
4350	Standby, calling clearance delivery.
4355	Say again call sign.
4400	Sage 26?
4405	Sage 36, roger, standby, expect lengthy delay.
4510	Sage 80. Willy clearance delivery, standby clearance on request, expect lengthy delay.
4517	Willy 80, hold in chalks.
4524	Roger.
5132	Reno 4, clearance, we're still waiting your clearance from Phoenix.
5555	Reno 4, clearance, you're cleared to Williams airport, Reno 2, maintain flight level 270, squawk 4366.
1:0910	Calling clearance, say again call sign.
1:0915	Willy 80, you're cleared to Williams airport, Clark 7, maintain flight level 190, departure frequency will be 317.5, sqauwk 4222.
1:0933	Readback correct.
1:0938	Standby.
1:0944	Say again call sign.
1:0946	Basket 1, roger, clearance on request.
1:1119	Basket 1, expect a delay in your clearance, base ops is putting you in the computer.
1:1249	Calling clearance say again.
1:1253	Basket 1, roger, base ops is putting your clearance in the computer, unknown delay.
1:1648	Sage 36, clearance delivery, you're cleared to Davis-Montham airport, on departure change route to read fly runway heading, at 3 DME turn left 210 radar vectors Stanfield, Stanfield 133 radial, Tucson 295 radial, then as filed, maintain 5 thousand, expect 15 thousand, 3 minutes after departure, departure frequency will be 363.0, sqauwk 4363.
1:2155	Roger, we're working on your clarance, standby.

1:2615 Falcon 18, Phoenix is trying to get your clearance out of the computer, now.
1:2707 Say again call sign.
1:2711 Reno 07, roger, clearance on request.
1:2757 Reno 07. clearance on request, standby.
1:2921 And Reno 07, expect a delay in your clearance, base ops is putting it in the computer.
1:2942 Willy 59, clearance cn request.
1:3155 Falcon 1, I have your clearance.
1:3159 Falcon 1, your're cleared to Chanlet Conn, correction, you're cleared to Williams airport, Chanlet Conn one, maintain flight level 190, departure frequency will be 317.5, sqauwk 4164.
1:3518 Willy 59, base ops is working on your clearance, standby.
1:3557 Reno 07, are you up.
1:3614 Reno 07, you're cleared to Williams airport, Reno 2, maintain flight level 270, departure frequency will be 317.5, squawk 4372.
1:3814 Reese 58, clearance on request, standby.
1:3831 Standby.
1:3915 And Reese 58, did you file through our base ops.
1:3922 Roger, expect a delay they're putting it in the computer.
1:3936 You filed at Reese, not here.
1:3958 And Reese 58, how do you spell your call sign.
1:4254 Reese 58, I have your clearance, ready to copy?
1:4259 Reese 58, your're cleared to the Reese airport, Superstition 9, Saint John transition, then as filed, maintain flight level 270, expect flight level 330, 10 minutes after dieparture, departure frequency will be 317.5, squawk 4120.
1:4325 Reese 58, you are cleared on the Superstition 9 departure, Saint John transition, then as filed.
1:4447 59, advice ready.
1:4453 Willy 59, you are cleared to IR 272 entry point, maintain VFR untii entry point, departure frequency will be 317.5, squawk 4243.
1:4513 4243, readback correct.
1:4921 Willy 15, standby, your clearance on request.
1:4941 15, I have your clearance, advice ready.
1:4945 Willy 15, clearance, you are cleared to Miramar airport, Stellar 5 departure then as filed, maintain flight level 190, expect flight level 350, squawk 4115.
1:5142 Standby, clearance on request.
1:5755 Calling clearance, say again call sign.
1:5800 17, roger.

2:0009 Willy 90, you're cleared to Williams airport, Shark 7, maintain flight level 190, departure frequency 317.5, squawk 4275.
2:0150 Stand, ah..., standby, I got clearance on request, emergency in progress.
2:0237 21, clearance.
2:0241 Willy 21, you're cleared to March airport, Stellar 5 as filed, maintain flight level 190, expect flight level 370, squawk 0741.
2:0755 19, standby, I have your clearance.

2:0926 Willy 19, I have your clearance, advice ready.

2:0932 19, you're cleared to March airport, Stellar 5 departure, then as filed, maintain flight level 190, expect flight level 350, squawk 0104.

2:1514 Willy 17, you're cleared to March airport, Stellar 5 departure, then as filed, maintain flight level 190, expect flight level 390, squawk 2627.

=====> 2:2000

Appendix C - Grammar Definition Files

```
;
;   Grammer Definition File for local controller voice recognition
;
.OBJ/ .DIGIT@2 .AGNT/ check-wheels-down/ taxi/ report/ cleared/ .LOC/ .HLD/

& without-delay/ .WIND/ .ALT/ .TRF/ change-to-departure/
roger

.OBJ=
      awol
      hook
      hungry
      spydr
      willy
      reno

.AGNT=
      tower
      willy tower

.HLD=
      hold
      hold-short
      and-hold .RSN/

.RSN=
      awaiting atc-release
      awaiting separation

.LOC=
      .DIGIT@2 .DIR/
        into-position/ across/ runway .DIGIT@2 .DIR/
      .DIGIT@1 miles .DIGIT@2 .DIR
        final-app-fix
      to-land
      option
      touch-and-go
      takeoff
      low-approach

.DIR=
      center
      left
      right
```

```
.WIND=
      wind calm
      wind .DIGIT@3 at .DIGIT@1,2

.ALT=
      altimeter .DIGIT@4

.TRF=
      traffic .ACF/
       traffic on-final
      traffic inbound
      number .DIGIT@1 to-follow-a .ACF

.ACF=
      T-38
      T-37
      F-5

.DIGIT=
      zero
      one
      two
      three
      four
      five
      six
      seven
      eight
      nine
```

```
;
;  Grammer Definition File for ground controller voice recognition
;

.OBJ/ .DIGIT@2 .AGNT/ continue/ taxi/ proceed/ .APP/ .HLD/ .LOC/ .WIND/
& .ALT/ .BIRD/ .RPT/ .CONT/

.OBJ=
     awol
     hook
     hungry
     spydr
     willy
     reno

.AGNT=
     ground
     willy ground

.APP=
     frequency-change/ back-taxi/ approved

.HLD=
     hold
     hold-short

.LOC=
     .DIGIT@2 .DIR/
     across/ to/ runway .DIGIT@2 .DIR/

.DIR=
     center
     left
     right

.WIND=
     wind calm
     wind .DIGIT@3 at .DIGIT@1,2

.ALT=
     altimeter .DIGIT@4

.BIRD=
     bird-condition .CLR/

.CLR=
     yellow
```

```
green
     red

.RPT=
       report-back
       report-backup

.CONT=
      contact clearance-delivery .DIGIT@3 point .DIGIT@1

.DIGIT=
     zero
     one
     two
     three
     four
     five
     six
     seven
     eight
     nine
```

```
;
;   Grammer Definition File for clearance delivery voice recognition
;

.OBJ/ .DIGIT@2 .AGNT/ cleared/ clearance/ .DEST/ .DEP/ .TRANS/ .DCT/
& .FILE/ .MTN/ .EXP/ .TIME/ .FREQ/ .SQW/

roger

.OBJ=
    awol
    hook
    hungry
    spydr
    willy
    reno

.AGNT=
    clearance
    clearance-delivery
    willy clearance-delivery

.DEST=
    davis-montham airport
    fort-huachuca airport
    williams airport
    norton airport

.DEP=
    .NAME .DIGIT@1 departure

.NAME=
    clark
    kilo
    shark
    reno
    stellar
    peter
    superstition

.TRANS=
    drake transition
    st-john transition

.DCT=
    needles
```

```
.FILE=
        then-as filed
        as filed
        available
        on request

.MTN=
        maintain .DIGIT@1,2 thousand
        maintain flight-level .DIGIT@3

.EXP=
        expect .DIGIT@1,2 thousand
        expect flight-level .DIGIT@3
        expect .DIGIT@1,2 minute delay
        expect lengthy delay

.TIME=
        .DIGIT@1,2 minutes after departure

.FREQ=
        departure-freq .DIGIT@3 point/ .DIGIT@1

.SQW=
        sqauwk .DIGIT@4

.DIGIT=
        zero
        one
        two
        three
        four
        five
        six
        seven
        eight
        nine
```

Appendix D - Prolog Display Program

Prolog program to display recognizer output:

```
predicates
        start
        process(char)
        repeat
        reader

goal start.

clauses
        start if
                makewindow(1,7,7,"Vocalink S4000 - Prolog Com1 Program",0,0,25,80),
                write("Prolog Com1 Program - Press <RETURN> to start."),
                readln(F),
                write(F),
                readdevice(com1),
                reader,
                readdevice(keyboard).

        repeat.
        repeat if repeat.

        reader if
                repeat,
                readchar(X),
                process(X),
                fail.

        process('\0') if write('\n'),fail.
        process('#') if !.
        process(X) if write(X),fail.
```

Biographical Sketch

Captain Robert F. Hall received his Bachelor of Science degree from the United States Air Force Academy in 1978. He is attending Arizona State University under the Air Force Institute of Technology, Civilian Institute program. Before attending ASU, Captain Hall was an Airlift Control Element Operations Officer and a C-5A Aircraft Commander at Travis AFB., California. He acquired over 2000 C-5A pilot flying hours. His additional duties included work in squadron and wing computer systems.

END
DATE
FILMED
DTIC
9-88