Natural Language Processing for Unmanned Aerial Vehicle Guidance Interfaces

by

Emily M. Craparo
Bachelor of Science, Aeronautics and Astronautics,
Massachusetts Institute of Technology, 2002

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of
Master of Science in Aeronautics and Astronautics
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2004

© Massachusetts Institute of Technology 2004. All rights reserved.

Author . . . .
Department of Aeronautics and Astronautics
May 30, 2004

Certified by . . . .
Eric Feron
Associate Professor
Department of Aeronautics and Astronautics
Thesis Supervisor

Certified by . . . .
Robert C. Berwick
Professor
Department of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by . . . .
Edward M. Greitzer
H.N. Slater Professor of Aeronautics and Astronautics
Chair, Committee on Graduate Students

Natural Language Processing for Unmanned Aerial Vehicle Guidance Interfaces

by

Emily M. Craparo
Bachelor of Science, Aeronautics and Astronautics, Massachusetts Institute of Technology, 2002

Submitted to the Department of Aeronautics and Astronautics
on May 30, 2004, in partial fulfillment of the
requirements for the degree of
Master of Science in Aeronautics and Astronautics

Abstract

In this thesis, I address the opportunities and challenges involved in applying natural language processing techniques to the control of unmanned aerial vehicles (UAVs). The problem of controlling an unmanned aircraft via natural language inputs is formulated as a feedback control problem, and notions of state, controllability, and observability are defined under this model. An implementation of such a system is also described. The phraseology of the existing air traffic control language is used as a base command set; to form this command set, a corpus of air traffic control commands was gathered from recorded exchanges between pilots and controllers at Boston's Logan Airport, as well as Hanscom Field in Bedford, MA, and these were used as the "target language" for this implementation. Knowledge of air traffic control and airport operations, combined with existing natural language processing techniques, is used to achieve a higher recognition success rate than a traditional natural language processor designed for a more general domain of discourse would. This is the first known attempt at formalizing air traffic control phraseology for use in an unmanned system.

Thesis Supervisor: Eric Feron
Title: Associate Professor
Department of Aeronautics and Astronautics

Thesis Supervisor: Robert C. Berwick
Title: Professor
Department of Electrical Engineering and Computer Science

Acknowledgments

Far from being a solitary endeavor, research thrives and generally depends on a successful networking of many individuals. During the last two years I have had the good fortune to come into contact with many remarkable individuals, a number of whom were instrumental in my research efforts.

I owe a great debt of gratitude to my advisor, Eric Feron. Aside from providing inspiration, ideas and direction for my research, he has also done a great deal to improve the quality of my graduate education, as well as my quality of life. His flair for finding unique opportunities for his students and for putting them into contact with the right people is remarkable, and his flexibility and willingness to accommodate requests have enabled me to have a hand in shaping the direction of my graduate education. His persistent support and encouragement have gotten me through some hard times. Most of all, his creative energy and passion for innovation have shaped my notion of research as it should be, and have provided much motivation for my work here.

Perhaps my most valuable source of technical information and assistance has been Robert Berwick. His wealth of knowledge in many areas of language processing, as well as his resourcefulness and practical expertise have greatly facilitated my acquisition of knowledge of language processing. He has also shown me great generosity, in terms of both his time and his resources, and has helped me immeasurably.

Missy Cummings has also been an outstanding source of information to me. Besides sharing her first-hand experience in military aviation and her unique insight into communication and information transfer from the perspective of a Naval aviator as well as an expert in human-computer interaction, she also put me into contact with other valuable sources of information on military aviation communication and gave me the opportunity to discuss the ramifications of my research directly with those who may one day utilize it.

A word of thanks is also due to my flight instructor, CFI John Nutt. Aside from "merely" doing a fantastic job of teaching me to fly, he has also been a principal source of information on air traffic control, continuing to clearly, thoroughly, and cheerfully answer my questions even after the completion of my flight training.

I would also like to thank Felix Chang and Jae Lee for providing me with the opportunity for technical collaboration in the early stages of this project's implementation. I very much enjoyed working with them, both for their great technical insight and their amiable personalities. I wish them both the best of luck in their future research endeavors.

Angela Copyak has also been an invaluable source of assistance for me. I have come to learn over the past two years that there is much more to research than its technical aspects, and Angela has greatly facilitated many of the vital processes necessary for my work to be accomplished efficiently. Not only has she done an excellent job as an assistant to Eric Feron, but her friendly and caring nature has also made our department a much more pleasant place to work.

Finally, I would like to thank my parents for all of their love and support over the years. They have always done their best to provide guidance, encouragement, and a sense of pride. Through their example I have learned the value of work ethic, balance, and integrity. These qualities, perhaps more than anything else, have enabled me to accomplish many of my goals in research and in life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Background

This thesis is a first attempt to explore the possibility of applying existing natural language processing techniques to the domain of air traffic control. A formal framework for the incorporation of natural language processing into a feedback control loop is developed, an implementation of a simple natural language processing system is described, and questions for future research are posed.

The remainder of this chapter gives a brief review of the fields of unmanned aviation, natural language processing, and air traffic control; and establishes motivation for the research presented in this thesis. Chapter 2 elaborates on the past and potential future application of natural language processing techniques to the air traffic control language. Chapter 3 gives an overview of a possible architecture for a language processing system for use in air traffic control, while Chapter 4 describes an implementation of the system outlined in Chapter 3. Chapter 5 discusses the behavior of this system in the context of the theoretical constructs set forth in Chapter 3. Finally, Chapter 6 describes some of the remaining open questions in this field, as well as possible directions for future continuation of this research.

1.1 Unmanned Aerial Vehicles in Perspective

1.1.1 Beginnings

Aerodynamic powered flight by unmanned aircraft is almost as old as human flight itself, but not nearly as well-developed. Although the first unmanned aircraft were flying successfully by 1918 [41], it was years before technological developments in the fields of automatic stabilization, remote control, and autonomous navigation allowed the widespread use of remotely controlled aircraft in the 1930s and the first all-target drone Navy squadron, Utility Squadron Five, which was established in 1941 [41]. Unmanned aircraft, in the form of German cruise missiles, saw their first military combat situations in 1944 [41]. Thirty-eight percent of these cruise missiles, almost four thousand units, were destroyed by British defenses, while the Allies suffered the loss of nearly three thousand aircrew members defending against these attacks [41]. Thus, the Germans demonstrated the utility of UAVs in reducing loss of human life. The 1950s saw the advent of reconnaissance drones, as well as significant advances in inertial navigation technology, largely by Charles Stark Draper of the Massachusetts Institute of Technology [28]. The use of UAVs for target prosecution began in the 1960s, with the Navy's Gyrodyne QH-50 [41]. The next major challenge facing UAV developers was that of increasing the vehicles' autonomous capabilities – a challenge that still stands today.

1.1.2 Current and Near-Term Uses

Military

> *I was looking at Predator [imagery displays] yesterday... It was flying over an area...at 25,000 feet. It had been up there for a long time, many hours, and you could see the city below, and you could focus in on the city, you could see a building, focus on a building, you could see a window, focus on a window. You could put a cursor around it and [get] the GPS latitude and longitude very accurately, remotely via satellite. And if you passed that information to an F-16 or an F-15 at 30,000 feet, and that pilot can simply put in that latitude and longitude into his bomb fire control system, then that bomb can be dropped quite accurately onto that target, maybe very close to that window, or, if it's a precision weapon, perhaps it could be put through the window...I'd buy a lot of UAVs in the future.*
>
> Admiral William A. Owens
> Vice Chairman of the Joint Chiefs of Staff
> June, 1995

The last few decades have seen a significant increase in the number of unmanned aerial vehicles (UAVs) in operation, and an even larger proliferation of interest in the supporting ideas and research concepts relating to UAVs. Much of this recent research and development has been performed with military applications in mind; Department of Defense budget allocations for UAVs have risen greatly in the last decade, with some programs more than doubling their funding from the 2003 fiscal year to the 2004 fiscal year [12]. The Department of Defense is projected to increase its level of investment in UAVs threefold in the first decade of the twenty-first century, resulting in the spending of over $10 billion on UAVs in that decade [4] (see Figure 1-1). UAVs also have Congressional support – in the 2003 fiscal year, Congress accepted all Department of Defense budget requests for UAVs, and in some cases added funding beyond that which was requested [4].

Recently, UAVs have been successfully deployed by Army, Navy and Marine Corps units in Operation Desert Storm and Operation Desert Shield [67], and one of the largest operational users of UAVs, the United States Air Force, utilized over ten types of UAVs in Operation Iraqi Freedom [4] [65]. The Air Force Scientific Advisory Board has noted the success of these vehicles both in traditional intelligence and combat situations such as surveillance, reconnaissance, target acquisition, target prosecution, command and control, meteorological data acquisition, medical resupply, battle damage assessment, and nuclear, chemical and biological agent detection;

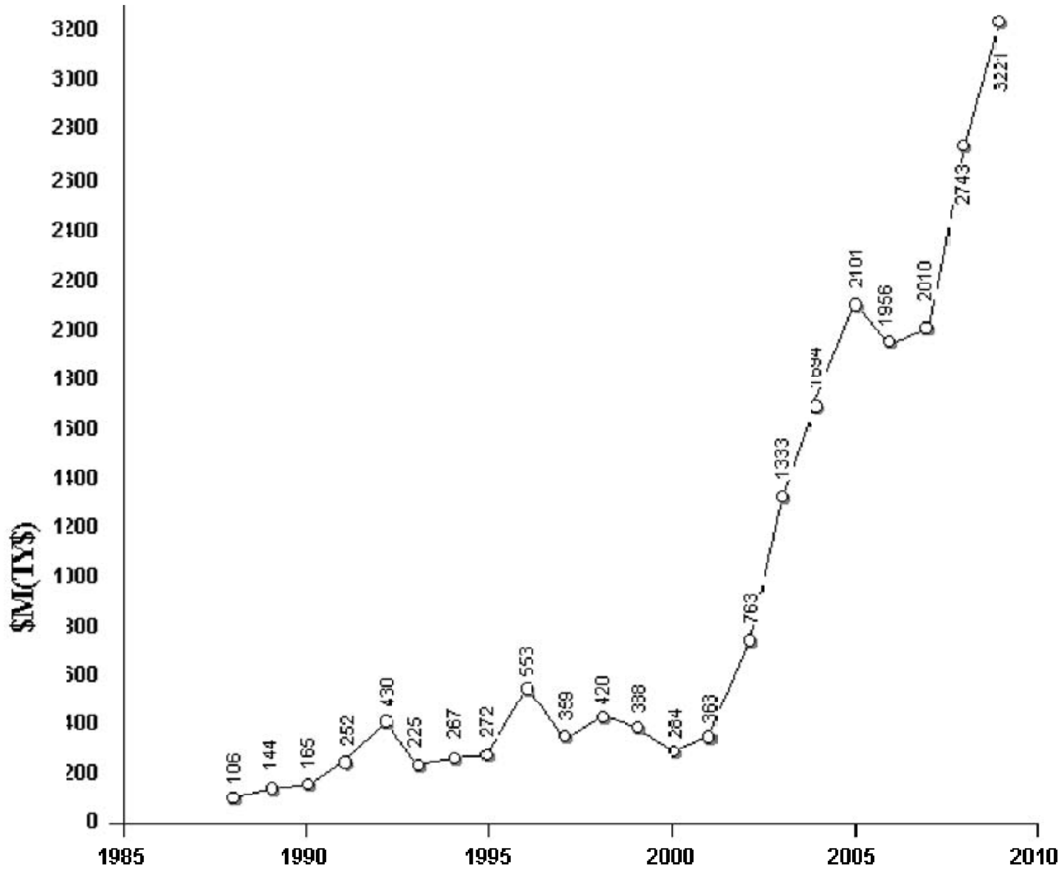


Figure 1-1: Past and projected Department of Defense funding for unmanned aerial vehicles.

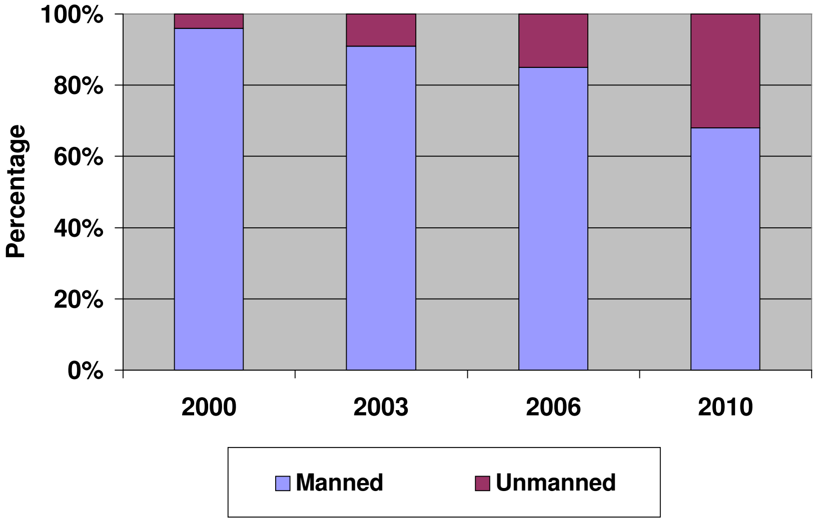


Figure 1-2: A shift in balance: Relative spending on manned versus unmanned aircraft. It is predicted that aircraft funding for unmanned aerial vehicles will increase from 4% in 2000 to 31% in 2010.

as well as in more novel missions such as those undertaken by the Special Forces [44] [65] [67]. Many military missions, particularly sentry-type operations, benefit from the extended loiter time that UAVs allow, and the obvious benefit of the reduction in risk to human life in such dangerous missions as suppression of enemy air defenses (SEAD) brought about by UAVs is particularly appealing [44]. Recently UAVs have expanded their repertoire beyond unarmed operations and into strike missions; armed remotely piloted vehicles have successfully engaged targets in the recent conflict in the Middle East, and a Predator UAV engaged in the first unmanned air-to-air combat incident with a manned Iraqi fighter aircraft in March, 2003 [4].

As unmanned aircraft have proven their worth in combat situations, and as recent advances in enabling technologies have allowed truly sophisticated unmanned aircraft to become more feasible, more and more emphasis has been placed on these vehicles relative to their manned counterparts, as shown in Figure 1-2 [4].

This fact was recognized in 1994 by the United States Air Force Scientific Advisory Board (USAF SAB), which undertook a comprehensive study of current and future needs of the Air Force, aimed at identifying the emerging technologies that would prove most pivotal and useful in the coming decades. Their report, entitled *New World Vistas: Air and Space Power for the 21st Century*, recognized the importance of unmanned vehicles in future military endeavors. Here, it discusses the merits of unmanned combat aerial vehicles (UCAVs) [66]:

> The UCAV is enabled by information technologies, but it enables the use of aircraft and weapon technologies that cannot be used in an aircraft that contains a human. There will be missions during the next three decades that will benefit from having a human present, but for many missions the uninhabited aircraft will provide capabilities far superior to those of its inhabited cousins.

Moreover, in addition to ably fulfilling many missions in their own right, UAVs are also effective "force multipliers;" that is, they increase the effectiveness of other units, generally by improving situational awareness or by providing cover or acting as decoys in dangerous situations [31] [44]. As the USAF Scientific Advisory Board noted following its examination of the performance of UAVs in Afghanistan and Iraq, "the benefit of Predator as ISR support for the C-130 gunship underscores the positive synergy between manned and unmanned systems" [65]. The Army has also made plans to make use of this "positive synergy;" a directive issued by the Department of Defense in 2003 mandated that future purchases of such Objective Force helicopters as the RAH-66 Comanche helicopter include a companion UAV [1] [4]. The Army also envisions UAVs as being a tool to provide live video feed to the crew of the Apache [25]. Further integration of unmanned systems with manned systems aimed at further exploitation of such benefits is thus a key area for future research.

The USAF Scientific Advisory Board has recently identified several key missions in which UAVs may have particular effectiveness, and has outlined the enabling technologies that will need to be developed to support these missions (see Figure 1-3) [65].

Indeed, the variety of missions withing the grasp of UAVs, as well as their remarkable effectiveness in carrying out these missions, has been recognized at the highest levels of our leadership [5]:

> [The Predator UAV] is able to circle over enemy forces, gather intelligence, transmit information instantly back to commanders, then fire on targets with extreme accuracy. Before the war, the Predator had skeptics, because it did not fit the old ways. Now it is clear the military does not have enough unmanned vehicles. We're entering an era in which unmanned vehicles of all kinds will take on greater importance – in space, on land, in the air, and at sea.
>
> President George W. Bush, December 2001

Civil

There are also many civil applications for which UAVs may prove useful. As the military has discovered, UAVs' typical advantage over manned aircraft in terms of

---

[1] The Comanche program was canceled entirely in early 2004, amid claims that the armed reconnaissance missions performed by Comanches could be handled entirely by UAVs [46]

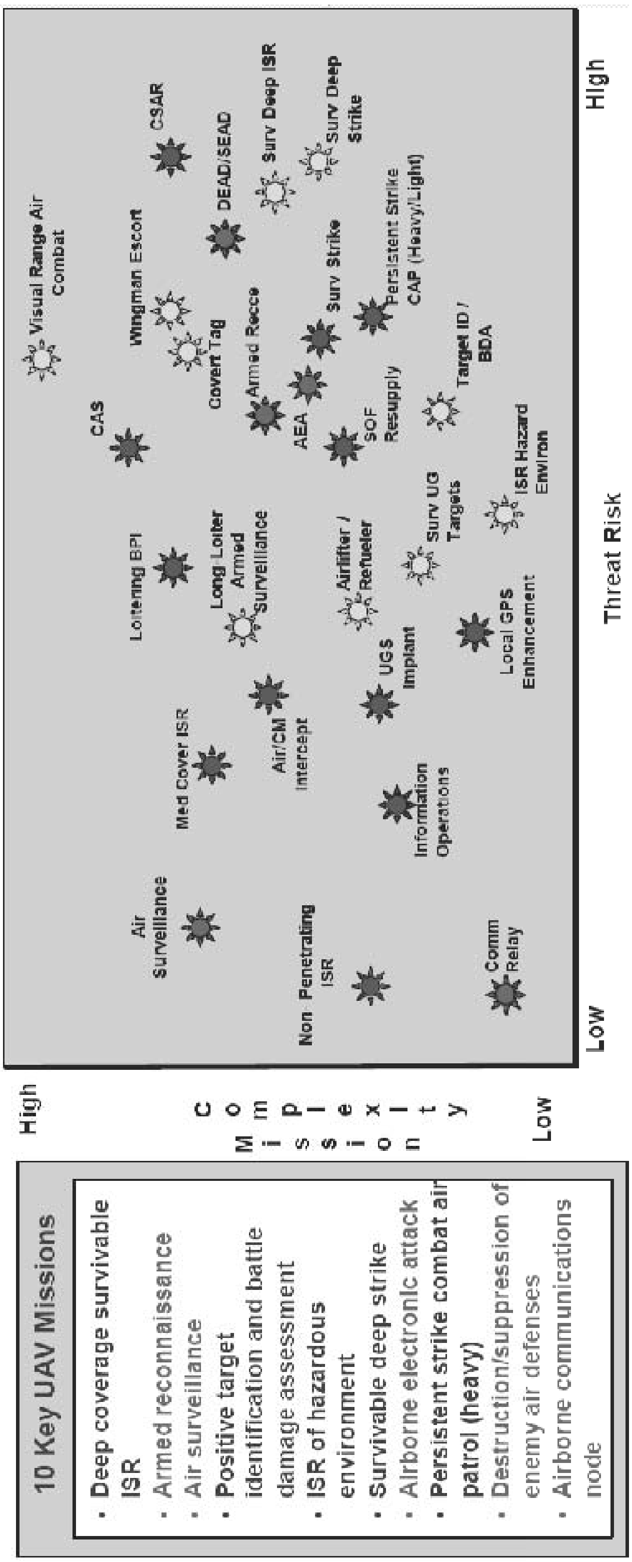


Figure 1-3: Key missions for unmanned aerial vehicles as identified by the United States Air Force Scientific Advisory Board.

loiter time makes them ideal for surveillance applications. Domestically, this capability may also be exploited for homeland security matters such as border patrol. As of early 2003, over 10,400 border patrol agents were assigned to the southern border of the United States, not to mention the various physical barriers, cameras, sensors, vehicles and manned aircraft that are deployed to combat illegal border penetration [3]. Despite this heavy and costly surveillance effort, the border is still vulnerable to penetration by human traffickers, drug smugglers, arms dealers and terrorists [3]. If the great success of UAVs in aiding military surveillance efforts is any indicator, border patrol operations could be made much more effective and perhaps less costly through the use of UAVs. This benefit is particularly relevant in an age of mounting concerns over acts of terrorism on American soil. Thus, the U.S. Border Patrol currently has plans to deploy UAVs, while the Coast Guard intends to use UAVs to monitor the nation's coastal waters [59]. Other possible uses of UAVs by the Department of Homeland Security and the Department of Transportation include such activities as monitoring the nation's major freeways and following trucks with hazardous cargo, as well as monitoring oil, gas, and power lines for purposes of protection and maintenance [4] [59].

UAVs may also prove useful to law enforcement. Currently, law enforcement agencies spend approximately $150 million annually for around 250,000 manned flight hours, much of which could be accomplished with UAVs [52]. UAVs have been proposed for such uses as drug interdiction and in aiding police in the pursuit of suspects, and could replace manned aircraft already used for law enforcement surveillance such as that undertaken during the October 2002 sniper incident in Washington, D.C. [4].

Other government agencies may utilize UAVs for such applications as fire detection and monitoring, control and monitoring of wildlife, disaster management, communication relays in the wake of natural disasters, inspections of buildings and bridges, and traffic surveillance [55] [52]. The Department of Energy has announced plans to test UAVs capable of detecting potential nuclear reactor accidents [3]. Innumerable lives could be saved with improved search and rescue operations, and UAVs could be a key element in providing more rapid exploration of unknown and especially inhospitable environments.

Scientific

Accompanying the recent increase in military and civil use of UAVs has been an increased awareness of their potential uses in the scientific community. UAVs are already recognized as useful and relatively cost-effective platforms for the collection of ecological, oceanographic, climatic, meteorological, and even volcanic data [52]. In fact, they often hold an advantage over manned systems in these areas, given their capabilities for long-duration flight and the fact that they can conduct scientific research in areas that are inaccessible or inhospitable to humans; for example, they can be flown into dangerous storms in order to collect more data. A NASA-sponsored program has produced and tested UAVs capable of monitoring pollution and measuring ozone [3], and the Department of Energy established a program in 1993 to study clouds and climate change; currently, the Department of Energy is investigating the possibility of using UAVs to measure radiation in the atmosphere [4] [59].

Commercial

Finally, UAVs have found advocates in the private sector as well. Besides their recreational uses, UAVs have the capability of furthering various commercial enterprises. Industries that may benefit from relatively inexpensive aerial photography include the real estate industry and the entertainment and film making industries [40]. For example, aerial surveys of properties and farmlands are often needed, and improved coverage of long-distance sporting events such as the Paris-Dakar race may be possible. The agriculture industry could certainly be made more efficient through the use of UAVs, and experiments have already been performed to explore the potential of UAVs in monitoring crops in order to determine optimal harvesting or watering schedules [60]. Solar-powered aircraft capable of being operated continuously for months at a time are being investigated as an enhancement or alternative to current broadband communication infrastructures [52]. Finally, commercial fishing operators currently rely on costly and short-duration manned flights for fish reconnaissance, and could benefit from incorporation unmanned vehicles (which could even perform shipboard takeoffs and landings) into their operations [52].

The recent expansion in both demand for and availability of UAVs for military, civil, scientific and commercial missions has brought to light the necessity for improved UAV control schemes. In particular, the need to have UAVs and manned aircraft operating safely and easily in the same shared airspace will become more pronounced as the number of UAVs in operation increases and their mission profiles become more intertwined with those of manned aircraft. Thus, it will become necessary to centrally coordinate both the manned and unmanned aircraft in a structured, yet intuitive fashion. While military interests have historically dominated the UAV industry, it may be civil and scientific entities that ultimately lead the way in including UAVs in the domestic national airspace; thus far it has been the Coast Guard that has spearheaded the effort to accommodate UAVs in civil airspace, with NASA and the commercial UAV industry also contributing to this effort [4].

1.2 Natural Language Processing

Just as UAVs are becoming increasingly ubiquitous, natural language processing (NLP) is making its way into more and more areas of everyday life. Despite the century-long pursuit of linguistics as a scientific discipline and the near half-century history of computational linguistics, it has only been in the last decade that practical language processing and understanding systems have entered the marketplace. Current applications of natural language processing include automated telephone customer service aids with speech processing capabilities such as AT&T's "How May I Help You?" system [24]; a variety of translation and language tools available via search engine sites such as Google and AltaVista; automated weather reporting programs that accept raw data and pass information on to the public in plain, unedited English [7]; and grammar and style checkers in commercially available word processing software.

Although language processing has been slow to appear in many applications due to the immense technical challenges associated with it, it is nevertheless one of the most comfortable and efficient ways in which a human can interact with a computer. Indeed, it is not surprising that this should be so, for people already regard computers as social entities. For example, studies show that a human is more apt to rate a computer's performance positively if the computer has recently said something flattering about the human. Humans are also more likely to rate a computer's performance positively if that computer asks for the rating than if a different computer asks the same questions [53]. Given this predisposition toward treating computers as one would treat other people, it is only fitting that human-computer interfaces mirror existing human communication systems.

While the idea of verbal communication with machines may seem exotic or far-fetched, it is nonetheless recognized as an important emerging technology by computer industry leaders. Microsoft founder Bill Gates has attested [21] to the appeal that natural language interfaces hold in envisioned systems:

> ...our long-term approach has worked very well for us, whether it's building standards or doing software innovation. We bet the company on MS-DOS. We bet the company on graphical interface. We bet the company on Windows NT. Now we're betting the company on these natural interface technologies, that that will bring computing to the next level of pervasiveness.

1.2.1 Voice and Language Processing in Aviation

Although little has been done to apply natural language processing techniques in the field of aviation, a number of new aviation systems are beginning to utilize speech recognition technology. It is widely believed that enabling voice control of various pieces of aviation automation will increase bring about a more efficient and less error-prone mode of operation. For instance, the Air Force has recently begun testing of voice recognition software for use in its Airborne Warning and Control System (AWACS) E-3 Sentry aircraft [39]. This technology would enable air battle managers to control their radar screens using voice commands rather than mouse, keyboard, and function key inputs. Preliminary testing has indicated that speech recognition software may enable battle managers to achieve a 40% reduction in workload, with increase accuracy and situational awareness [39]. As shown in Figures 1-4, 1-5, 1-6 and 1-7, experiments show that a consistent decrease in mental workload, physical demand, effort ratings, and necessary communication events may be realized uniformly in all phases of AWACS operation through the use of voice commands [70].

In addition to these measurable improvements in performance when using voice commands, AWACS controllers also had highly favorable reactions to the voice in-

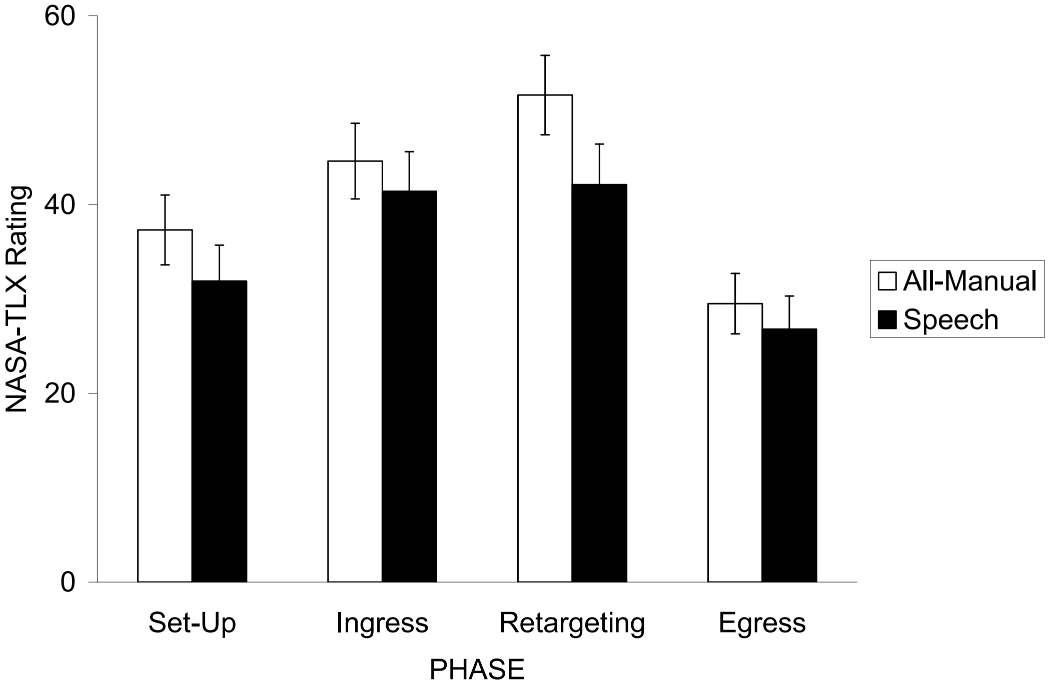


Figure 1-4: Mean NASA-TLX ratings and standard errors of mental workload of AWACS controllers as a function of phase and control modality.

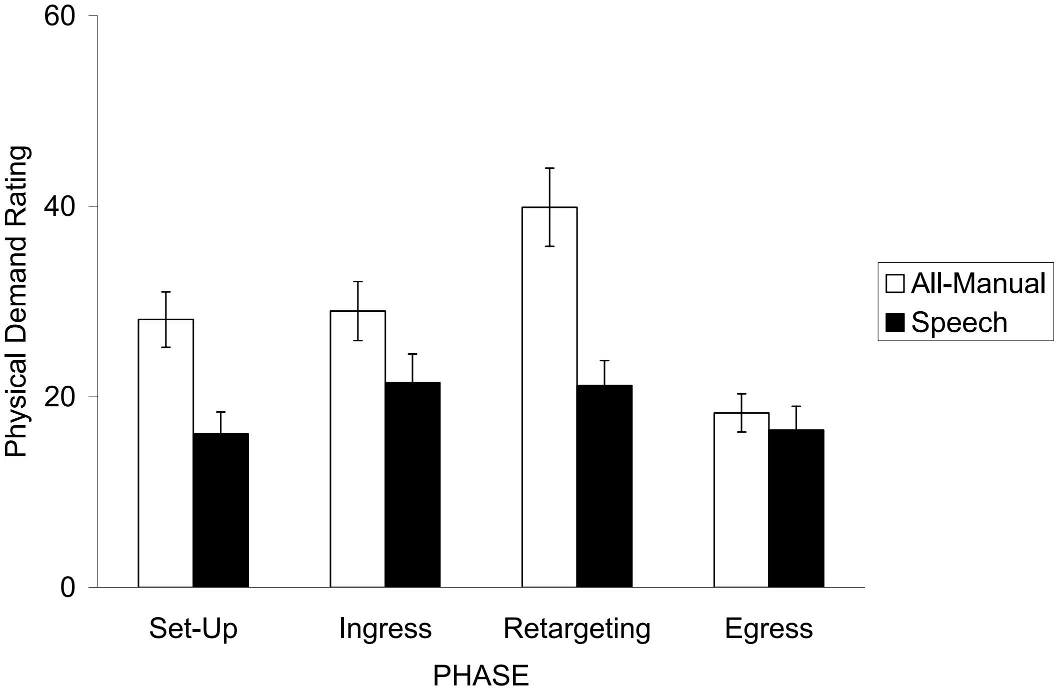


Figure 1-5: Mean Physical Demand ratings and standard errors as a function of phase and control modality.

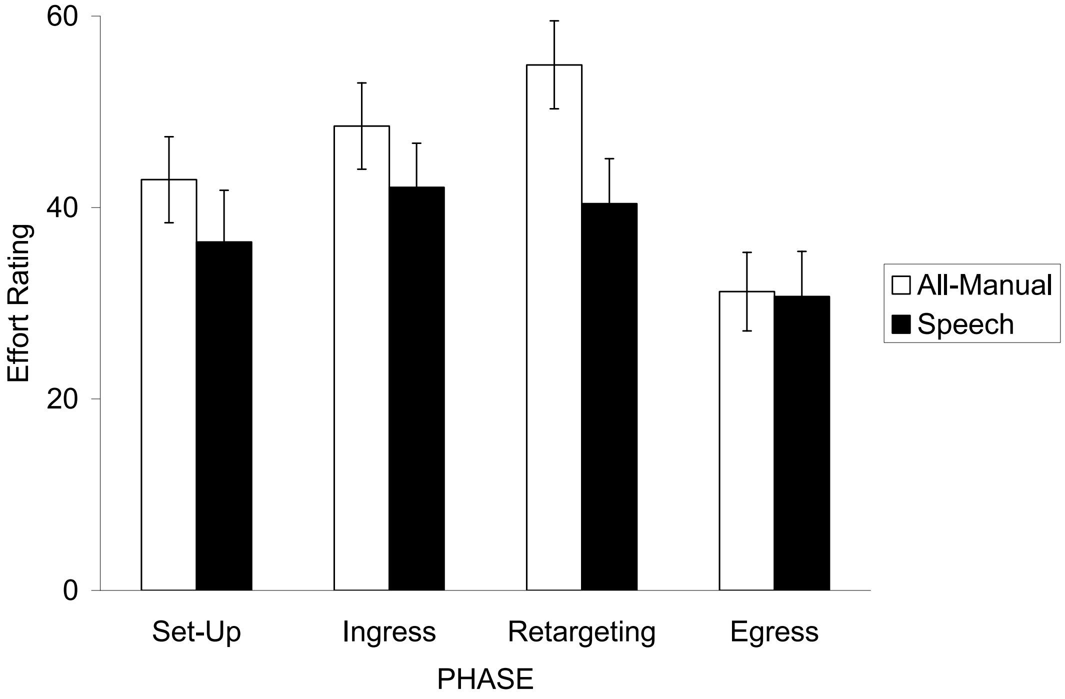


Figure 1-6: Mean Effort ratings and standard errors as a function of phase and control modality.

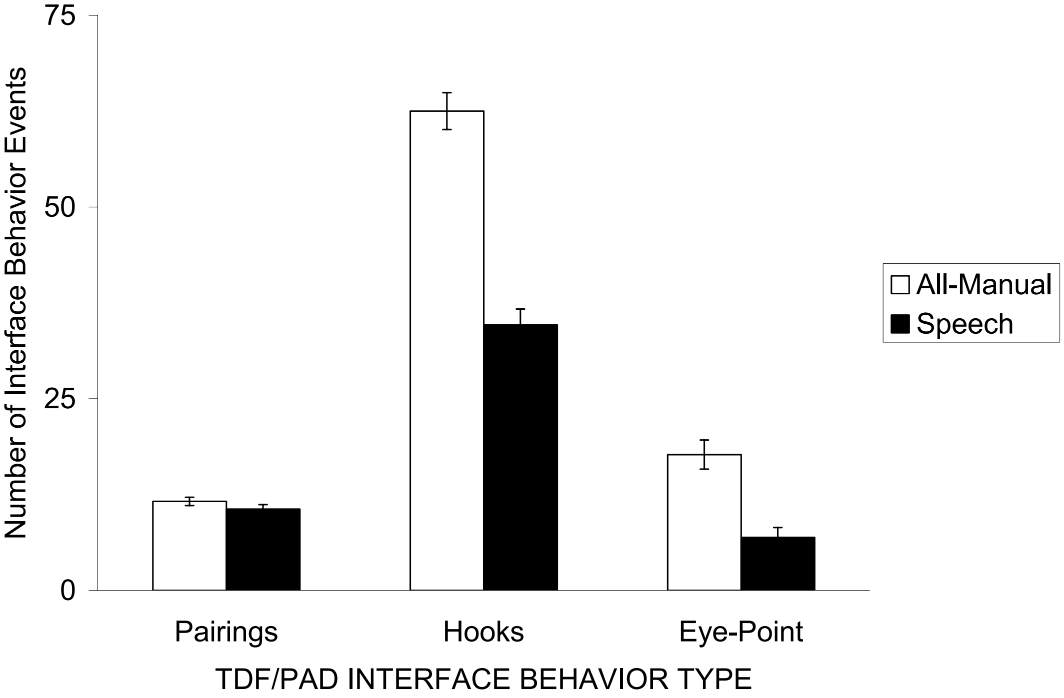


Figure 1-7: Mean number and standard errors of three TDF/PAD interface activity measures as a function of control modality.

terface and indicated that they felt comfortable and at ease with the interface [70]. Many seemed to appreciate the decrease in workload and the quickness with which tasks could be accomplished via the voice interface, with one respondent noting that

> The speech controls would be an excellent addition to future AWACS interfaces. Given the fact that future command and control will be even more focused on time sensitive tasks, speech controls will allow future [air weapons officers] the ability to handle more complex and a greater number of tasks in the same amount of time.

These positive sentiments were echoed by most participants in the experiment; as Table 1.1 indicates, a large percentage of participants preferred to utilize voice commands when both voice and manual commands were available to them, and as Table 1.2 shows, the majority of participants agreed that voice control was either "very useful" or "extremely useful" in most tasks, and all agreed that it had at least "minor utility" in all tasks [70].

Similar technologies to the voice-activated AWACS workstations have been proposed for use in "interactive kneeboards" intended for use by Navy fighter pilots, with the goal of achieving a comparable reduction in workload and "head-down time" and an increase in accuracy [47].

TASK	*%*
Set-up Phase, Open ATO	50
Set-up Phase, Mark Controller	100
Set-up Phase, Sort ATO	64
Set-up Phase, Set Bulls-eye	91
Ingress Phase, Hook Aircraft (Tag)	100
Ingress Phase, Check-in Aircraft	67
Ingress Phase, Open ATO	18
Ingress Phase, Hook Aircraft (Pair)	82
Retargeting Phase, Repairing	100
Switch Bulls-eye Probe Task	83
Range & Bearing Probe Task	58
ATO Question Probe Task	42
Threat Call Hooks	10

Table 1.1: Percentage of participants using speech to accomplish speech-enabled tasks during preference trials.

Utility Rating Question	*Waste of Effort*	*Minor Utility*	*Somewhat Useful*	*Very Useful*	*Extremely Useful*
Interacting with ATO	0	1	3	7	1
Interacting with TDF/PAD situation display	0	0	9	3	0
Responding to bearing & range query	0	2	2	6	1
Pairing friendly tracks against targets	0	1	0	3	8

Table 1.2: Number of participants that selected each response category for four utility rating questions on a post-experimental speech interface survey form.

1.3 Air Traffic Control

The modern system of air traffic control is a marvel of logistical organization, responsible for assuring the safe and efficient transgress of approximately 50,000 civil and military aircraft per day in United States airspace and up to 5,000 aircraft per hour during peak travel times [20]. The remainder of this section describes the history of air traffic control [42], as well as some possible directions for its future development [72].

1.3.1 Evolution and Current System

Modern aviation could not exist without the air traffic control system, but early aviators in the United States had no such agency responsible for maintaining separation and providing direction. Due to technological limitations of early aircraft, virtually no flight took place at night or during unfavorable weather conditions. Thus, the practice of "see and be seen" was generally an adequate collision avoidance measure for early pilots. Although the establishment and rapid expansion of transcontinental airmail service and commercial airline services during the 1920s greatly increased the amount of traffic in the nation's skies and prompted President Calvin Coolidge to begin regulating the use of airways in the mid-1920s, the slow-flying aircraft of the era and the requirement that pilots remain clear of clouds and fly only where the visibility was at least three miles remained adequate countermeasures against aircraft collisions.

By the 1930s, however, instrumentation and navigation technology had advanced far enough to allow pilots to fly at night and in marginal weather conditions. While this improved technology allowed pilots to navigate more safely and reliably in a wider range of conditions, it also allowed them to get into situations in which they would not be able to "see and avoid" other aircraft. This problem was most acute in the vicinity of the ad hoc sod fields that served as the airports of the day, due to both the relative congestion of these areas and the lack of standard operational procedures and, in some cases, runways.

It was to remedy these near-airport difficulties that the first rudimentary air traffic control system was established. The first air traffic controller, former barnstormer Archie League, was employed by the St. Louis Lambert Municipal Airport in 1929 (see Figure 1-8). Stationed on a beach chair near the approach end of the field and equipped with a pair of colored flags, League used his flags to signal pilots with instructions to either hold or continue in for landing or takeoff. Although this simplistic system lacked the sophistication and flexibility of today's system, it provided a passable improvement in safety and airport efficiency at the time, and other major airports soon hired air traffic controllers of their own.

Improvements in air traffic control soon came with the construction of raised, enclosed control towers and the replacement of colored flags by colored light guns. Using the light guns, air traffic controllers were better equipped to communicate with aircraft at night an in poor weather conditions, could more effectively direct their messages at a particular aircraft, and could eventually construct more meaningful

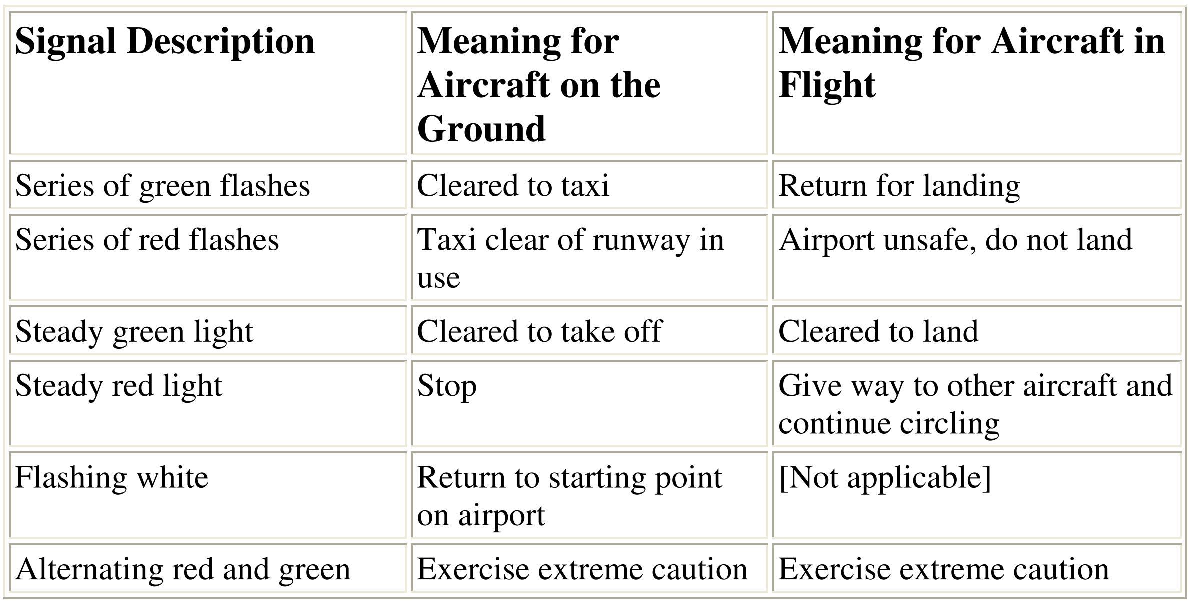
Figure 1-8: Archie League, the first air traffic controller.

messages from the three colors of light available to them than they could with only two flags (see Table 1.3).

It is worth noting that the light gun system of air traffic control still exists today in a form not very different from that utilized in the 1930s. Although it is currently reserved for use with aircraft lacking radios or in the event of a radio malfunction, the fact that such a simple set of signals is still considered useful for directing air traffic is significant.

While the light gun system allowed improved communication from air traffic controllers to pilots, it was still impossible for pilots to initiate communication with controllers. This problem was rectified with the introduction of radio communication equipment into air traffic control towers. Early air traffic control radios allowed controllers to communicate with pilots of radio-equipped aircraft up to 15 miles away from the airport, in virtually all weather conditions. Radio communication allowed controllers to give pilots any information they wished, including sequencing information and weather and airport advisories, and pilots were able to respond to controllers and even initiate communication. Moreover, because pilots did not have to be looking at the airport control tower in order to receive a message, they were able to concentrate on navigation and traffic monitoring without risking a missed communication. Thus, the introduction of radio communication greatly increased airport safety and efficiency.

However, there still existed flaws in the air traffic control system. Few airlines



Signal Description	Meaning for Aircraft on the Ground	Meaning for Aircraft in Flight
Series of green flashes	Cleared to taxi	Return for landing
Series of red flashes	Taxi clear of runway in use	Airport unsafe, do not land
Steady green light	Cleared to take off	Cleared to land
Steady red light	Stop	Give way to other aircraft and continue circling
Flashing white	Return to starting point on airport	[Not applicable]
Alternating red and green	Exercise extreme caution	Exercise extreme caution

Table 1.3: Modern light gun signals.

were willing to install the heavy, expensive radio equipment on their aircraft, and small aircraft were often incapable of generating the electrical power required to run it. Additionally, although radio communication allowed for a much greater range of information to be passed from controllers to pilots, there were still no standards governing the type of information passed or manner of its transfer. Pilots initiated contact with the control tower at widely varying distances away from the airport, or sometimes not at all, since pilots were not required by law to contact air traffic controllers at that time. Because there was not yet a standard phraseology in use, many pilots did not understand the instructions being given to them.

The first steps toward federal regulation of air traffic began with the creation of the Bureau of Air Commerce in 1934. The Bureau established rules to govern flight in instrument meteorological conditions, and requested that the major airlines take steps to ensure aircraft separation on congested flight routes between airports. The airlines responded by establishing airway traffic control units (ATCUs), which regulated aircraft flying under instrument flight rules (IFR traffic) on major airways and provided some guidance for aircraft flying under visual flight rules (VFR traffic). This system was successful in reducing the number of near misses that occurred, but it still lacked a set of standardized operational procedures for separation of aircraft and handing over of control to airport controllers, and pilots were still under no legal obligation to file flight plans or to participate in the air traffic control system.

Many important changes to air traffic control were made during World War II. By 1943 aviation had become the largest industry in the world, and roughly 85 percent of IFR flights in the U.S. were military flights. In order to ameliorate the overly heavy burden placed on control towers by the increasingly congested airways, newly established approach control facilities began to handle separation of traffic 15 to 20 miles away from the airport. Additionally, early versions of today's flight service stations were established to give preflight weather briefings and file flight plans. A bitter power struggle took place during the early 1940s between civilian and military air traffic control authorities, the result of which was that existing civilian facilities would continue to take responsibility for separation of both civilian and military aircraft, although the Army did establish some control facilities of its own for military aircraft.

In another step toward international standardization of air traffic control, the International Civil Aviation Organization (ICAO) was established in 1947. This organization set the U.S. systems of navigation and communication to be the standard for the world, and also declared English to be the common international language of air traffic control.

During the 1950s and 1960s, air traffic control was plagued by high controller attrition, an overstretched budget, and mounting numbers of increasingly fast and high flying airliners that demanded more and more airspace between aircraft in order to maintain separation standards. Several high-profile accidents took place in the late 1950s and early 1960s, including the midair collision of two airliners over New York City [42] (see Figure 1-9 [33]). These accidents, along with other factors, prompted leaders to begin an overall examination and reconstructing of the national air traffic control system. The Federal Aviation Agency (FAA), a precursor to the modern Federal Aviation Administration, was created in 1957 as an independent agency responsible for the administration of the air traffic control system and the oversight of government research and development in aviation. An increased effort was also made to incorporate radar surveillance equipment into air traffic control facilities. In 1961, President Kennedy instructed the FAA to "conduct a scientific, engineering overview of our aviation facilities and related research and development and to prepare a practicable long-range plan to insure efficient and safe control of all air traffic within the United States." In response, the FAA created Project Beacon, a task force responsible for investigating and evaluating the air traffic control system. The task force's report called for increased attention to the immediate problems facing air traffic controllers, such as inadequate radar and transponder equipment and obsolete information transfer techniques. This resulted in the implementation of the flight data processing (FDP) system, as well as improved radar technologies and displays.

During the late 1960s, it became apparent that aviation would need to be more fully integrated with other forms of transportation in order to reach its full potential. In an effort to promote this integration and centralization of efforts, the Department of Transportation (DOT) and the National Transportation and Safety Board (NTSB) were created, and the Federal Aviation Administration became a part of the Department of Transportation rather than an independent entity. Although this change resulted in a weakening of the power of the FAA, it helped to harmonize the future development of aviation and ground facilities.

The Airline Deregulation Act of 1978 resulted in record amounts of air traffic from myriad new and old airlines, as well as a general overhaul in air traffic route structures as many airlines moved toward the "hub and spoke" system commonly used today. Unfortunately, these large changes in air traffic activity happened at

Figure 1-9: The aftermath of a midair collision of two airliners over New York City in 1960. Although an investigation revealed pilot error as the cause of the crash, it also pointed out that more sophisticated surveillance equipment would have enabled controllers to detect the problem and issue corrective instructions to the pilots, thus hastening the installation of radar equipment at busy airports.

a rate too quick for the FAA to properly respond, and in many cases neither the equipment nor the number of controllers present at an airport was adequate for peak traffic flows. Worse still, illegal strikes by discontented controllers resulted in the firing of over 10,000 controllers in the early 1980s, exacerbating the situation. The ensuing shortage of controllers was in part offset by the implementation of novel flow control techniques by the FAA, but both delays and controller workload still increased at many major airports.

Throughout the organizational and bureaucratic metamorphosis undergone by the air traffic control infrastructure, its operational language remains relatively unchanged in the last several decades. This consistency reflects the high degree of appropriateness of the language for its task, as well as the degree to which the air traffic control language has evolved beyond most of the inconsistencies and ambiguities present in the English language. Thus, while air traffic control is a natural language that has evolved over time, it is also among the most highly formalized in existence.

Other such specialized languages have evolved for the purpose of time-critical guidance and control; examples include the phraseology utilized by surgeons when speaking to their assistants during surgery, and other transportation-related languages.

1.3.2 Future Changes

As the air traffic control system continues to evolve, it is likely to rely increasingly heavily on automation, both in the form of decision aids for controllers and pilots and improved communication systems between controllers and pilots. In recent years the capacity of U.S. airspace as steadily approached the maximum level possible under the current framework. At the same time, the chronic shortage of controllers that has historically afflicted air traffic control has not disappeared. Thus, some automated systems are intended to reduce controller and pilot workload in order to increase the efficiency of the system. As these automated systems are developed, one of them primary areas of concern for developers is the degree to which the systems can be easily and naturally interacted with by their operators. As we stand at the threshold of an integrated manned-unmanned system, it is vital that the needs of both humans and machines be addressed, all within the time-critical framework of the air traffic control system.

Chapter 2

Natural Language Processing in Feedback Control

2.1 Suitability of Air Traffic Control for Natural Language Processing

Current interfaces between humans and UAVs do little to recognize the human affinity for verbal communication or the accepted practice of guiding aircraft through oral commands. Existing control schemes, such as pure data link and radio control, greatly restrict the variety of commands that may be easily passed to a UAV. Radio control relies on continual monitoring and deliverance of low level commands, and it ignores the necessity of centralized planning in aircraft guidance. Data link, on the other hand, enables centralized control but also necessitates high controller workload and is inconsistent with current air traffic control practices at many airports. Both of these control schemes severely restrict the extent to which UAVs can be successfully integrated into the existing civil and military air traffic control system. This need not be, however - both the highly structured nature of air traffic control phraseology and the algorithmic and goal-driven nature of flight make air traffic control an ideal venue for the application and development of natural language processing technology.

Natural language interfaces provide a promising alternative to traditional UAV guidance and control interfaces. Such an interface would be more intuitive than conventional RC controls and would therefore require less training to operate. It would also reduce human workload because once commands are given, the machine would not need to be supervised while the commands are being carried out. Because a natural language interface can be made to utilize the existing air traffic control phraseology, a formal and structured yet intuitive framework for communication with vehicles, it is well suited to integration into the current air traffic control system. Finally, a computerized system would even offer certain advantages over a human RC pilot: theoretically, it should be able to perform its task optimally, not just safely, and would therefore yield increased efficiency.

2.1.1 Alternative Means of Communication: Data Link

In examining the possibility of utilizing a voice-oriented UAV guidance interface, it is useful to examine them in the context of comparison with another major emergent technology that has been proposed as a UAV guidance interface, as well as an aid to human pilots or possibly a replacement for verbal air traffic commands altogether: data link.

The data link system has been proposed to alleviate some of the shortcomings of the voice-oriented air traffic control system. It has been shown that 80% of information transfer problems in aviation occur over radio channels [72] and it is estimated that communication-induced delays result in the loss of over $300,000,000 annually for airlines [72]. These problems stem from factors such as overcrowded communications channels, as well as limitations on short-term memory and phenomena such as expectancy that may cause a communication to be misinterpreted (although this phenomenon occurs in visual perception as well). It is argued that with the semi-permanent visual display of information available, pilots will be less prone to forgetting what was said to them and will also be able to refer back to the command if in doubt of its content. Additionally, some information, such as navigation points, may be easier to perceive visually rather than through auditory channels.

While the data link system undoubtedly has benefits, there are some areas in which voice-oriented systems may hold benefits over data link systems for both controllers and pilots. The National Research Council has noted [72] that

> with regard to the workload of the data link task itself, there is considerable consensus that the composition and initiation of lengthy keystroke messages by either ground or air personnel involve considerably higher workload than spoken messages.

This higher workload translates into an increased amount of time required to send a message via data link. There is also another source of slowdown inherent in the use of data link systems, and that is the increased lag between the time a message is received and the time when a pilot or controller responds to it. Studies have shown that there is a 50-100% increase in the time it takes a pilot to initiate a response to a data link message versus a voice message, and similar increases for controllers [34] [71] [62] [17] [18].

In addition to the additional time required for a round of communication via data link, there is also an additional burden placed on the visual resources of both pilots and controllers. Pilots, in particular, must dedicate additional "head-down time" to read and cross-check data link messages, and studies show that they will do so even if it is their co-pilot's job to monitor and operate the data link [72] [49]. One proposed remedy to this problem is the addition of a redundant synthesized voice transmission, which has been demonstrated to reduce head-down time, and has the added benefit of eliciting a quicker response from pilots [22]. Indeed, combined voice-data link systems may hold considerable promise for increasing efficiency of air traffic control operations, both in terms of time and number of utterances – it has been shown that the total number of communications (voice or data link) is lower in a combined system than in a system that uses only voice or only data link [63] [64]

2.2 Previous Work

There has been very little previous work on natural language processing for UAV guidance interfaces, and even less work that utilizes the existing air traffic control phraseology.

The Stanford Computational Semantics Laboratory has investigated natural language guidance of unmanned aerial vehicles [13], [36]. Their work was aimed at multi-modal communication with UAVs, e.g. communication via voice and graphical inputs. While it addressed many important issues related to task-oriented dialog, it did not do so in the context of air traffic control. Similarly, researchers at Brigham Young University have developed voice- and PDA-based UAV control interfaces and have obtained good performance on voice command recognition [51], although their voice command set was also developed solely for their research and was not based on air traffic control commands, nor did it allow the kind of flexibility that the air traffic control language allows.

Some work has been done in applying natural language processing to air traffic control for use with manned systems, generally with the goal of reducing error due to miscommunication, ambiguous instructions, or errors of memory.

Cushing [11] has proposed a natural language processing system capable of acting as a "mediator" between air traffic controllers and pilots. Citing accidents due to ambiguity of commands and errors of interpretation, Cushing suggests passing communications through a processing unit capable of automatically filtering out potential sources linguistic confusion and asking the speaker for clarification if necessary. In addition to its applicability as an operational safeguard, Cushing notes, such a system might also be useful as a training device for pilots an controllers, alerting them to any potentially confusing verbal idiosyncrasies they may have. In addition, it could serve as one layer of a human-machine interface.

Churcher *et al* intended to use speech recognition technology to automatically transcribe certain, essential parts of transmissions between the air traffic controllers and airborne pilots [10]. They claimed that these transcripts could be used air traffic control training purposes, or for relaying information to the pilot in flight and thereby reducing pilot workload in a manner similar to data link. They used IBM's off-the-shelf commercial continuous speech recognizer, and it gave them only a modest accuracy of recognition (around 30%) in its base form. However, when the device was augmented with other knowledge sources and higher levels of linguistics such as contextual information and context-free syntax, the accuracy could be greatly improved to over 70% even in noisy environments. While this result is not spectacular, it does show a large improvement over the baseline result and indicates a promising area for future work.

An important example of such high-level knowledge is the structure of a discourse. Discourse is defined as a collocated, related group of sentences, and a group of sentences must be coherent in order to make a discourse [32]. There are at least two different approaches to coherence. One is an informational approach, where relationships between sentences impose constraints such as result, explanation, parallel, and elaboration on the information in the utterances [29]. Historically, this approach has been applied predominantly to monologues between a speaker and hearers. Another approach is called intentional approach, in which utterances are understood as actions, requiring that the listeners infer the underlying goal [27]. This intentional approach has been applied mostly to dialogs.

The notion of intentional coherence in discourse plays a significant role in the air traffic control system, because the high level tasks of landing, takeoff, and maneuvering around the airport must be coordinated by a group of sentences, not just individual sentences. Even human pilots are urged to take advantage of the predictable nature of flight discourse; a student guide to voice communications published by the United States Navy gives this advice [68]:

> Know what to **expect**. As you progress through each flight, you should know what is expected to happen. If you know what is to be said ahead of time, responding correctly will be much easier.

Thus, pilots are encouraged to rely on their knowledge of the established structure of both flight and communications to form their responses to ATC commands. The manual continues:

> Every conversation with a controlling agency or service follows a specific progression...proper communication involves the realization of which progression phase you are in and making correct and timely responses.

This aspect of communication is not lost on computational linguists. Grosz argued that a discourse could be represented as a composite of three interacting components: a linguistic structure, an intentional structure, and an attentional state [27]. Grosz also pointed out that task-oriented dialogs have a linguistic structure that closely parallels the intentional structure of the task being performed [26]. The fundamental notion in this observation is that a discourse has an underlying purpose which it aims to achieve, called the discourse purpose, and that each segment of a discourse also has a finer-grained purpose, called discourse segment purpose. Then they are organized in a tree-like hierarchy with two coherence relations: dominance and satisfaction-precedence. This structure helps a discourse management system understanding the intention of a speaker.

2.3 Current Research

Figure 2-1 illustrates the way in which a natural language processing module might fit into a feedback control loop: the human controller compares the UAV's current state with the desired state in much the same way as an air traffic controller. The controller then issues a command to the UAV, which is processed by the natural language

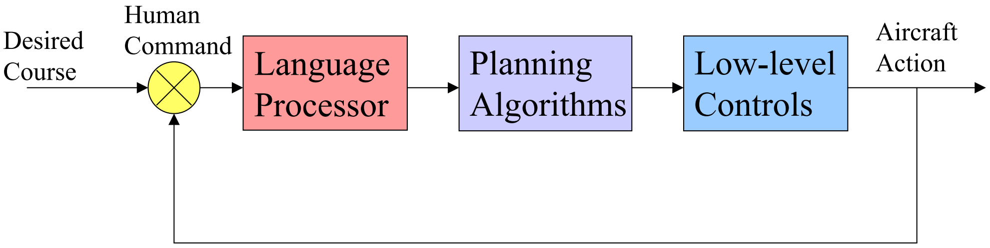


Figure 2-1: Incorporation of a natural language processor into a feedback loop.

processing module to produce a "machine language" command. This command may then be sent to a planning or optimization module [56], which then sends an exact trajectory into the low-level control system, from which appropriate commands are sent to the actuators and control surfaces of the aircraft.

In describing a feedback control system, there are various notions that need to be defined [69]. For example, the notion of stability is important for all feedback control systems. In the case of the guidance interface described here, the issue of controllability is of interest. In order for this system to be controllable, it must be possible to reach any desired state from an initial equilibrium state. Put another way, it must be possible to command the UAV such that it can perform any sequence of actions typically required of a manned system, within the UAV's own dynamic constraints. Observability of the system is also relevant in a guidance interface, and in this situation corresponds to a controller's knowledge of the system's interpretation of commands and intended responsive action to these commands. Given the time-sensitive nature of the system, it is also desirable that there always exist a control capable of delaying further state changes. In other words, the UAV should always be capable of "holding" in a safe configuration at any time, within its fuel constraints [57]. In this case, the desired output of the system might be taken to be the trajectory the controller wants the plane to fly (i.e. "fly heading 230"), or a certain condition he or she wishes to have met independently of the exact way in which it is met (i.e. "remain clear of the bravo airspace"), or the exchange of information (i.e. "what is your altitude?") [56]. Thus, it is necessary to define not just the state space of the aircraft, but also a database and a dynamically changing set of variables that describe the system and may be updated by the UAV's processors or the air traffic controller. Finally, in order that the performance of the system may be objectively evaluated, it is necessary to define performance metrics. Some possible metrics might be number of required communication utterances, time to acknowledge and complete an assigned task, user satisfaction and comfort, user workload, and, of course, safety.

Chapter 3

Overview of Example Implementation

The following two chapters describe the implementation of a basic natural language processing system built to interpret air traffic control commands and simulate appropriate aircraft responses. It is intended to be a first approximation to a fully operational-quality system; however, it addresses many challenges faced by dialog systems of every level of complexity, and despite its simplicity behaves well in many common situations encountered in modern air traffic control.

3.1 Design Considerations

A variety of carefully considered simplifications were made in order to define an appropriate scope for this implementation, while still retaining a certain degree of realism. The system is limited to near-airport operations, including ground operations, takeoff, and landing[1]. The initial state of the system is predefined, and a limited number of aircraft are allowed. Normal airport operating conditions are assumed.

The aim of this thesis is to examine the properties of the air traffic control language and the application of natural language processing techniques to it. Thus, the separate issue of speech recognition, which is an important topic in its own right, is not addressed in this thesis. Instead, text-based inputs are accepted, so that the focus of the project can remain rooted in language processing rather than speech processing.

Actual aviation communication involves information exchange among a wide variety of participants (see Figure 3-1 [75]). Even a brief flight under visual flight rules from a controlled airport will generally require multiple communication sessions with one or more controllers, as well as the independent acquisition of weather data from one or more sources. In this simulation, it is assumed that each aircraft will interact

---

[1] As described in Chapter 1, it was also in near-airport operations that the first air traffic control system was implemented. Just as the increased congestion around airports prompted the first efforts at controlling aircraft, the demand by both manned and unmanned aircraft for airport facilities will likely make this the first location in which they will need to be centrally coordinated.

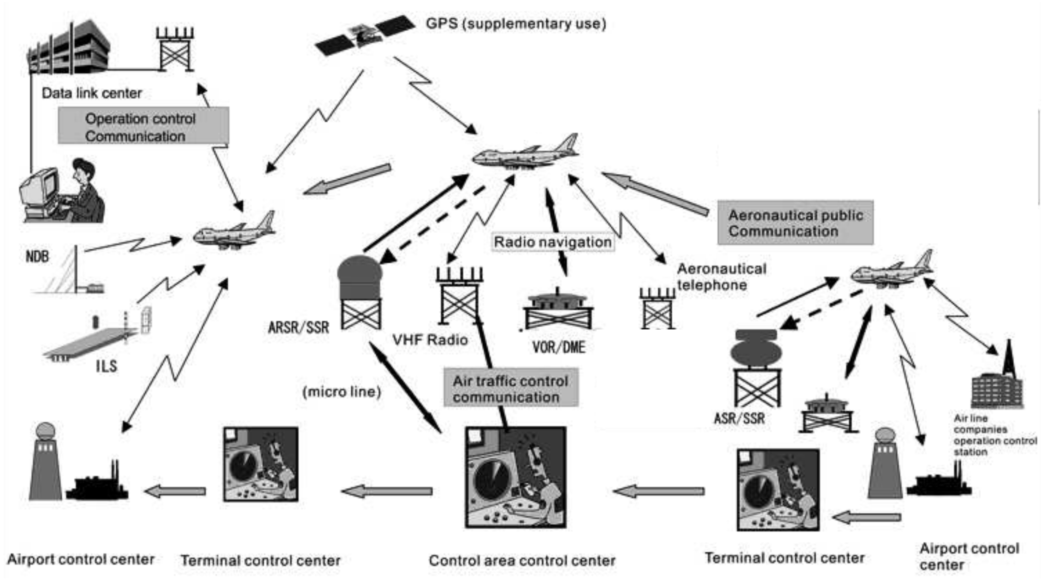


Figure 3-1: Schematic diagram of aviation communication.

only with a single controller. In a real system, necessary handoff procedures would need to be established, and independent or pilot-initiated communication procedures would have to be accounted for.

In order to assure controllability, the system was designed to accept "primitive" control commands consisting of the vectoring commands used in the modern air traffic control system (with which the author of this thesis has personal experience). Some examples of vectoring commands are "climb and maintain one five thousand," "reduce speed to two five zero knots," and "turn right heading two three zero."

While the primitive commands are sufficient to allow a controller to specify an exact trajectory for an aircraft, they are rather cumbersome and poorly suited to high-traffic situations. Thus, the system was built to understand a number of higher-level commands, in which the controller specifies a goal for the aircraft rather than an exact means to achieving this goal. An example of such a command is "taxi to runway two three," after which the aircraft is left to decide on its own how best to accomplish this.

System observability is attained in much the same way that controllers obtain their knowledge of pilots' intentions: Read back of commands. When a command is received and processed by the system, the system's assessment of what it should be doing next is read back to the controller. If the controller finds that his input has been misinterpreted, it can be canceled and a new input can be given. If, however, the system has understood the input correctly, no further confirmation is required from the controller.

3.2 System Overview

The system is comprised of five distinct modules as shown in Figure 3-2. A user playing the role of a ground and air traffic controller can issue text commands to the preprocessor, which then passes the edited commands to the sentence parser. The sentence parser recognizes a set of sentence structures, and converts sentences into standardized verb templates. These verb templates are then passed to the discourse manager. In the discourse manager, consecutive verb templates are analyzed for feasibility and inconsistencies. Inconsistencies are reported to the user in much the same way a that pilot would ask a controller for clarifications. If no inconsistencies are found, the verb templates generated by the discourse manager are transformed into primitive commands for the aircraft. For this simulation, a database maintains the current and immediate past states of the airport and the various aircraft in the system. It is consulted when the discourse manager checks for inconsistencies, and it is updated by the discourse manager when new commands arrive. Finally, a graphical simulator displays the airport and aircraft states, and it maintains the simulation clock.

The parser and dialog manager for this system were constructed based on a corpus of air traffic control transmissions collected from Boston's Logan International Airport and the Laurence G. Hanscom Field in Bedford, MA. This corpus contains data collected over a two-year period of time (2002-2003) and during a variety of typical operational situations, and contains the necessary utterances necessary to accomplish virtually all typical near-airport operations.

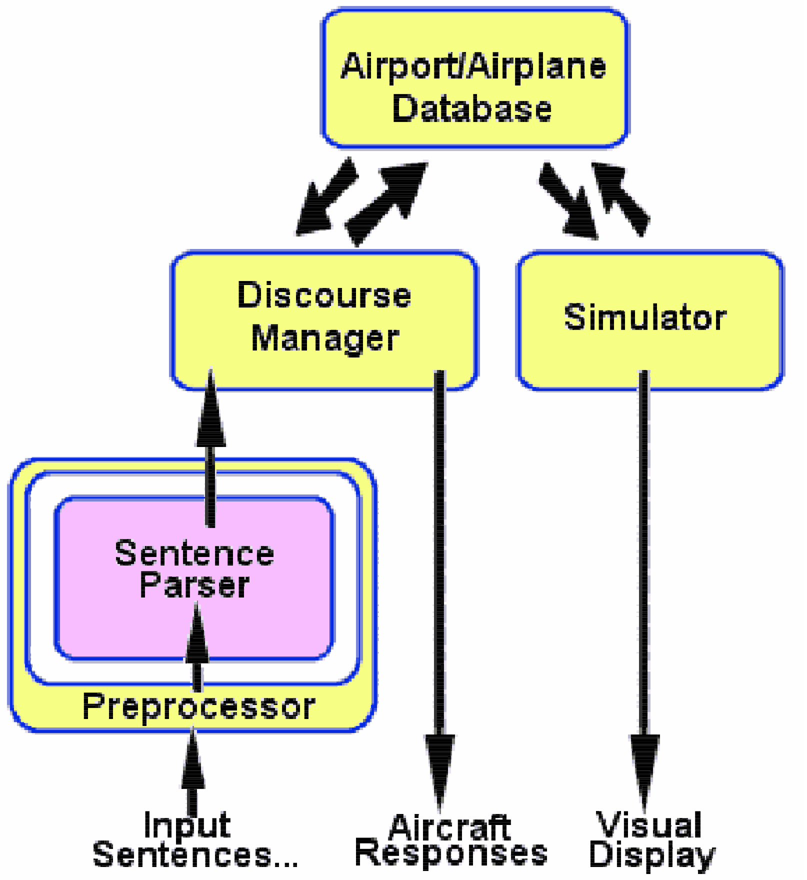


Figure 3-2: System layout.

Chapter 4

Detailed Description

4.1 Parser

Parsing is the recognition and assignment of structure to a string. Syntactic parsing consists of the recognition of sentences and the assignment of syntactic structure to them. Thus, in order to parse a sentence, each of its constituent parts must be given one or more labels, some formal syntactic structure for these labels must be defined, and the sentence must be processed in such a way that one or more parse trees are assigned to it.

4.1.1 Lexicon

Before a sentence may be assigned a particular structure, each of its constituent words must be identified and labeled. A lexicon provides a repository of words in a given language, and their relevant syntactic characteristics. A lexicon of words relevant to air traffic control would contain entries such as

$Number \longrightarrow one \mid two \mid three \mid ...$
$Determiner \longrightarrow a \mid an \mid the \mid ...$
$Noun \longrightarrow contact \mid runway \mid taxiway \mid ...$
$Verb \longrightarrow contact \mid taxi \mid turn \mid ...$
...

The lexicon for the system described in this thesis contains the verbs relevant to air traffic control, as well as new classes of words for airplane names, taxiway names, units of measurement, and numbers that conform to conventional air traffic control phraseology.

4.1.2 Grammar

In addition to a system for placing labels on each word in a sentence, it is also necessary to describe the ways in which these words are related to one another. A *grammar* is a set of rules describing the ways in which terminals (words) can be

Type	**Common Name**	**Rule Skeleton**	**Linguistic Example**
0	Turing Equivalent	$\alpha \longrightarrow \beta;\ \alpha \neq \epsilon$	Augmented Transition Networks
1	Context Sensitive	$\alpha A \beta \longrightarrow \alpha \gamma \beta;\ \gamma \neq \epsilon$	Tree-Adjoining Grammars
2	Context Free	$A \longrightarrow \gamma$	Phrase Structure Grammars
3	Regular	$A \longrightarrow xB$ or $A \longrightarrow x$	Finite State Automata

Table 4.1: The Chomsky Hierarchy. A and B denote a non-terminals, α, β, and γ denote strings of terminals and non-terminals (possibly empty, except where indicated), x denotes a string of terminals, and ϵ denotes the empty string.

represented by non-terminals (equivalence classes such as noun phrase (denoted in this thesis as NP), verb phrase (VP), and sentence (S)).

There are many possible grammatical frameworks, but the ones used most commonly by computational linguists are those found in the Chomsky hierarchy (see Table 4.1) [32]. Here, types of grammars are arranged in descending order of complexity. That is, a Type 0 grammar is capable of defining all languages defined by Types 1, 2, and 3, as well as some languages that cannot be defined by the other types of grammars, while Type 1 grammars can define all languages defined by Type 2 and Type 3 grammars, but not all languages defined by Type 0 grammars.

As Table 4.1 indicates, the various types of grammars are characterized by the forms of grammar rules that are allowed. Type 0 grammars are unrestricted; any non-empty string may be written as any other string. The class of languages defined by Type 0 grammars are the recursively enumerable languages, i.e. those that may be enumerated by a Turing machine.

In context-sensitive grammars, a non-terminal may be written as any non-empty string of terminals and non-terminals, provided it is in a certain context. It has been demonstrated that some natural languages, such as Swiss German, are context-sensitive [30] [58].

Context-free grammars allow a non-terminal to be written as a string of terminals and non-terminals, or possibly the empty string. Many natural languages appear to be described by context-free grammars, or at least closely approximated by them.

Finally, the regular grammars allow a non-terminal to be written as a string of terminals followed by at most one non-terminal. They are equivalent to regular expressions and so can be described by finite state automata.

From the point of view of someone wishing to formally describe a given language, the prospect of a finite-state grammar may be attractive. Finite-state grammars, however, are incapable of capturing some aspects of the English language. Chomsky [9] has shown that a language can be generated by a finite state automaton if and only if it can be generated by a context-free grammar that does not have any center-embedded recursions, i.e.

$$A \longrightarrow \alpha A B$$

For example, the following sentences may be constructed using center-embedded recursions:

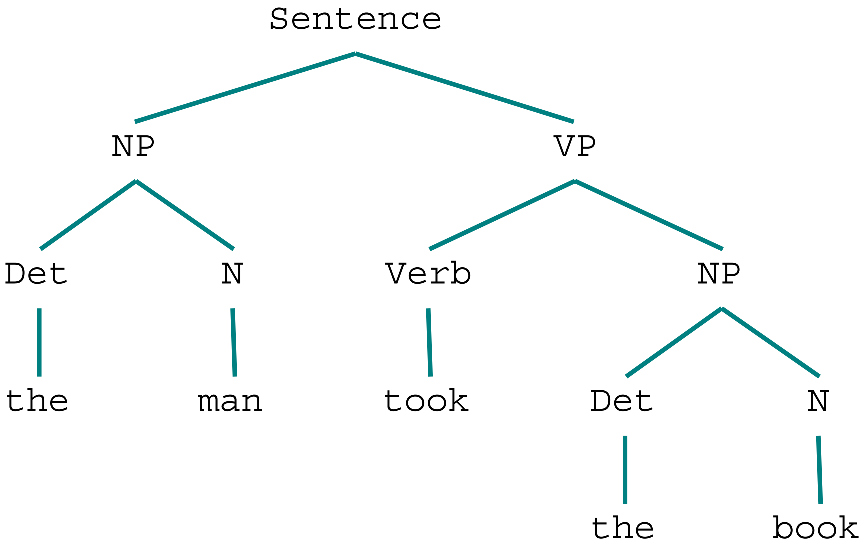


Figure 4-1: The first context-free grammar parse tree (Chomsky, 1956).

The man stole the car.
The man the policeman chased stole the car.
The man the policeman the woman called chased stole the car.
etc...

Another example involves sentences embedded in other sentences. Let S denote an English sentence. Then, another valid sentence is

The person who said S was incorrect.

This sentence may also be center-embedded within another sentence, and this process may continue indefinitely.

While difficult to read and understand[1], these sentences are perfectly grammatical according to the English syntax. Thus, while finite-state grammars can model many aspects of English and can often provide good approximations of it, not all English syntax can be modeled with a finite-state grammar.

It appears that English is, in fact, a context-free language [32]. No known analysis has been made of the air traffic control language. However, it is assumed that "aviation English," as a more constrained subset of English, is at most context-free as well.

A context-free grammar is formally defined as a 4-tuple [32] consisting of

1. A set of non-terminal symbols N

---

[1] It has been noted [50] that many of the constructions used to prove the complexity of grammars of natural languages also tend to cause severe difficulty in human comprehension. Thus, while it may not be possible to generate all *grammatical* English strings without resorting to a complex grammar, the subset of those strings that is *easily understood* by humans could possibly be generated by a simplified grammar, perhaps even a finite state grammar.

2. A set of terminal symbols Σ such that $\Sigma \cap N = \phi$
3. A set of productions P, where each is of the form $A \longrightarrow \alpha$, $A \in N$, $\alpha \in (\Sigma \cup N)*$
4. A start symbol S

In the implementation described in this thesis, N and part of Σ are defined by the lexicon, while P and the remainder of Σ are contained in the syntactic rules. S is implicitly contained in the syntactic rules as well.

4.1.3 Approaches to Parsing

Many algorithms have been developed for parsing strings according to context-free grammars. Such algorithms generally produce parse trees (such as the one shown in Figure 4-1) that depict the structure and organization of the input string. They do so by attempting to reconcile predefined grammatical and lexical constraints such as

$$\begin{array}{l} S \longrightarrow NP\ VP \\ NP \longrightarrow Det\ Noun \\ VP \longrightarrow VP\ NP \\ VP \longrightarrow Verb \end{array}$$

with the input string. Traditional approaches to parsing have included top-down parsing, in which all possible parse trees are generated breadth first until the leaves of one of them are consistent with the input, and bottom-up parsing, in which the input sentence is examined to see which rules it follows, and trees are constructed from the bottom up until one of them reaches an S node. While both of these approaches are intuitively easy to understand, they have shortcomings.

While the top-down parser does not waste time producing trees that cannot be rooted in an S node, it may produce many trees that are inconsistent with the input sentence. It is even possible for a top-down parser to run forever on certain grammars without ever completing a tree; see Figure 4-2. In addition, top-down parsers are also poorly equipped to handle ambiguous inputs. Since they only return the first tree they find, top-down parsers are bound to miss possible parses of sentences such as those shown in Figures 4-3. Finally, top-down parsers often repeatedly parse subtrees, discarding them when they do not cover the input and then parsing them again after backtracking. This leads to inefficiency.

Bottom-up parsers also suffer from extreme inefficiency; although they never produce trees that have no chance of matching the input, as top-down parsers do, they generate a large number of trees that never have a chance of reaching an appropriate start node.

The Earley Parser

The parser contained in this system is an extension of the Earley context free parser [15], which utilizes a dynamic programming [2] approach to eliminate the problems

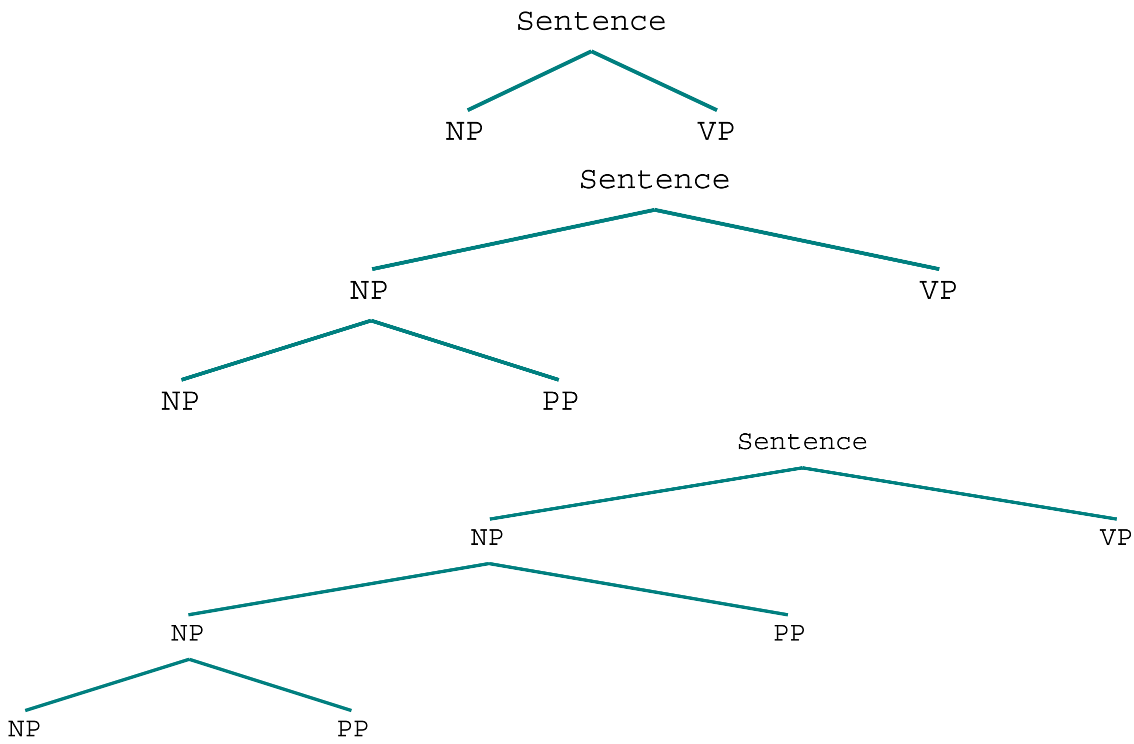


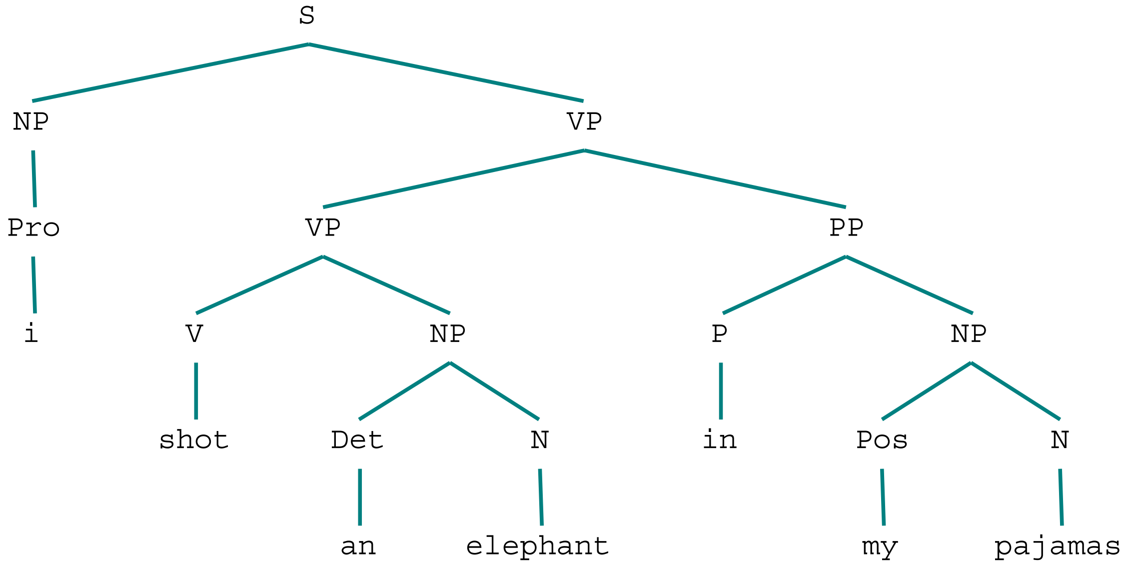
Figure 4-2: The beginning of an infinite expansion of left-recursive rules created by a top-down parser parsing left to right.

and efficiencies encountered by top-down or bottom-up parsers. In general, dynamic programming algorithms systematically solve and store the solutions to all sub-problems needed to solve the overall problem. In the case of parsing, this means that the various subtrees for all parts of the input are discovered once, and then looked up on subsequent reparses. Because there is a great deal of backtracking and regularity in the search space inherent in parsing, this can result in a significant gain in time efficiency. In addition, because all of the possible subtrees are stored in a single chart, all possible parses may be retrieved, including ambiguous ones.

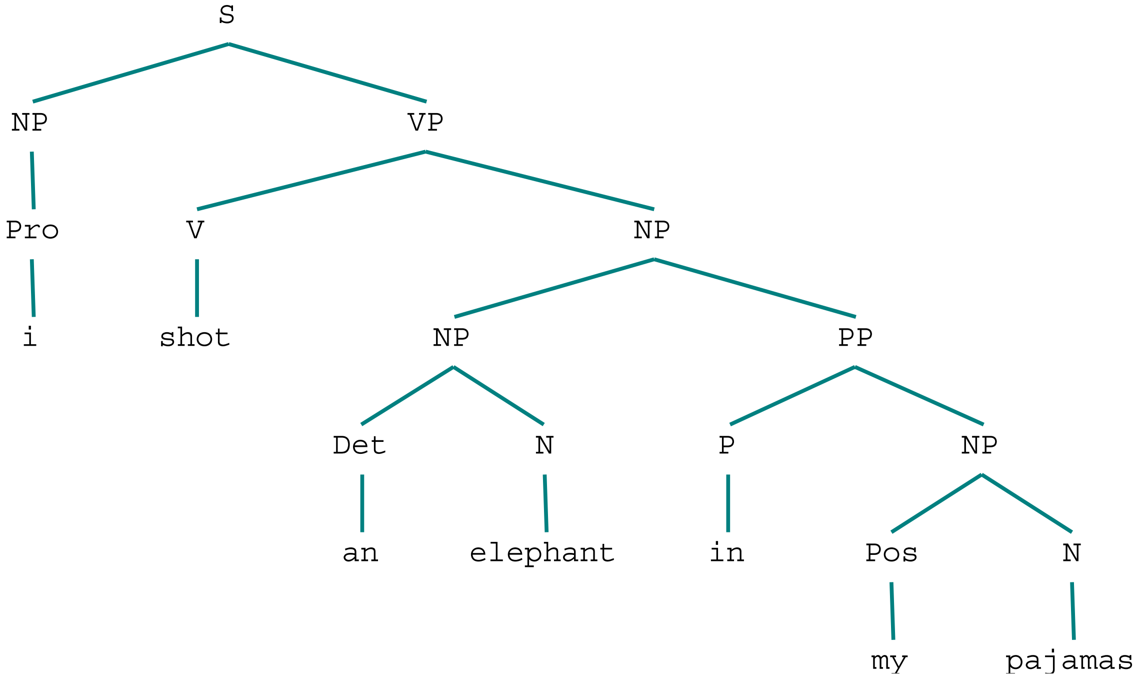
For this implementation, syntax rules were created to handle the rather truncated grammar of air traffic control, in which most sentences are imperatives and "unnecessary" words such as prepositions are often (but not always) omitted. The parser filters out semantically empty phrases such as "I want you to [*command*]" or "I'm going to have you [*command*]."

4.1.4 Verb Frames

In a goal-oriented language such as air traffic control, it is useful to formalize commands in terms of the action they are meant to induce, as well as any qualifications



(a) The semantic interpretation of this parse is that speaker shot an elephant while wearing pajamas.



(b) This parse indicates that the speaker shot an elephant that was in the speaker's pajamas.

Figure 4-3: Two valid parsings of the ambiguous sentence, "I shot an elephant in my pajamas."

that may exist on this action. For example, some command verbs, such as "contact," generally take one or two arguments (e.g. "contact ground" or "contact tower on 119.7"), while others may take an undetermined number of arguments (e.g. "taxi to runway five," "taxi to runway five via sierra," "taxi to runway five via sierra and echo," etc.). To capture this predicate-argument relationship between command verbs and their arguments, it is useful to employ a subcategorization frame of the type used in natural language processing to encode relationships between words and their complements [32]. While this project focuses on the relationships between command verbs and their arguments, it is also possible to create subcategorization frames for other parts of speech as well.

The verb frames generated by this system's parser are rather specialized - they output information with relevant headings based on the verb being parsed.

For example, the input sentence "taxi to runway three via zulu" generates

```
[OUTPUT](go :to (runway :num 3) :via (taxiway :num zulu) :agent
you)
```

while the sentence "turn three two zero" generates

```
[OUTPUT] (turn :heading (heading :to 320) :agent you)
```

This actor-action-object framework facilitates the interpretation of parser outputs by those parts of the system that must translate these outputs into actions.

4.2 Discourse Manager

Because the parser's context-free grammar rules were intentionally made as general as possible, many nonsensical sentences are parsed. For example, the same rules that would allow the parser to handle a sentence such as "follow that Continental to the runway" might also allow it to accept a sentence such as "follow that tower to the runway." Additionally, a sentence accepted by the parser may be semantically acceptable in isolation, but not when taken into context with surrounding sentences or the state of the aircraft. For example, the command "climb and maintain ten thousand" is semantically acceptable at the sentence level, but is inappropriate if the aircraft is already at an altitude of twelve thousand feet. It was to alleviate these problems that the discourse manager was created.

When the discourse manager (shown schematically in Figure 4-4) receives a verb template from the parser, it performs a semantic interpretation of the verb frame to determine what action to take. If it finds the requested action to be reasonable and unambiguous, it updates the states of aircraft in the system database accordingly and reads back this update to the controller. However, if the discourse manager finds the command to be nonsensical, inappropriate for the circumstances, or inconsistent with previous commands, it generates a response message requesting clarification or another command. The discourse manager takes the context of the input into account, and rejects a command requesting an action in conflict with the context even if the input is syntactically and semantically valid at the sentence level.

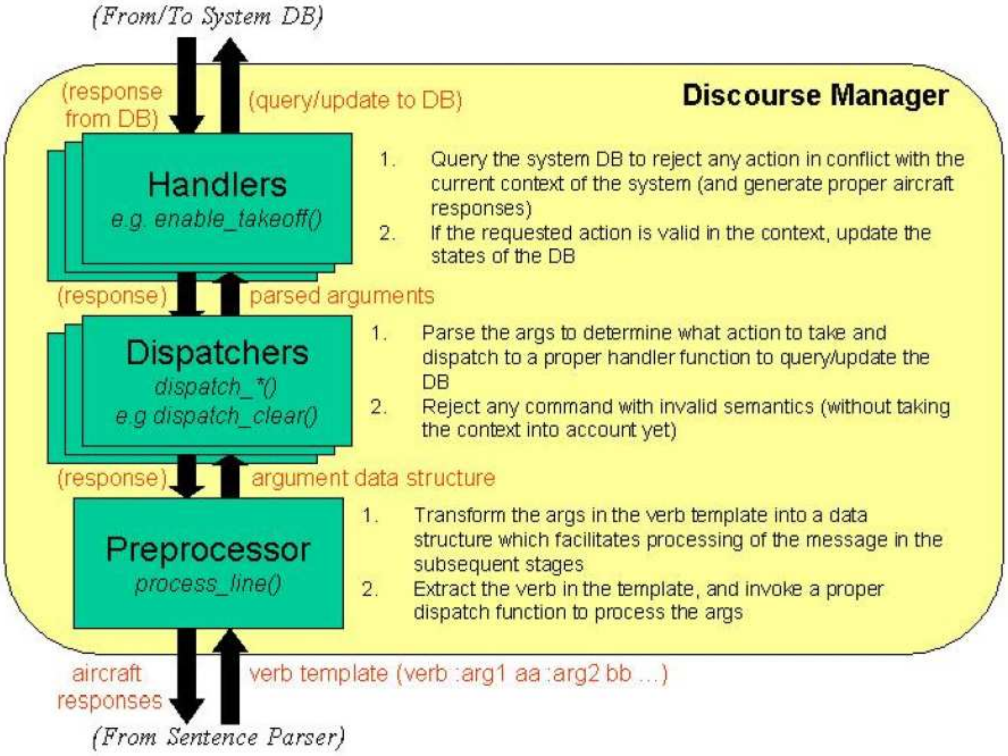


Figure 4-4: The discourse manager.

As shown in Figure 4-4, this module has three internal stages: a preprocessor, a set of dispatch functions, and a set of handler functions. The preprocessor transforms the arguments of a given verb template into a data structure containing (record, value) pairs, and invokes an appropriate dispatch function depending on the verb. Recall that the lexicon contains verb entries, such as “climb,” “descend,” “taxi,” “remain,” and “maintain.” Each verb entry has its own dispatch function in the form of dispatch_*(). For example, any command containing verb “maintain” is processed by dispatch_maintain().

It is the responsibility of these dispatch functions to further parse their arguments, and invoke the appropriate actions (such as changing altitude, or updating a value in the system’s database). At this stage, the dispatch functions can detect and reject messages that are semantically valid in the sentence level, but are not coherent in the greater context.

First in the dispatch functions’ analysis of a command is the notion of history. For example, there exist commands that nullify the effects of the previous command, such as “cancel that”, “never mind”, and “let’s not do that.” In order to resolve the meanings of “that”, it is necessary to keep track of the history of commands. In this system, it is assumed that only the immediately preceding command will be referenced, and so one-command history is kept. An example of the handling of a cancellation command is as follows:

```
TOWER: Delta, clear to taxi to two-niner.
```

```
PLANE: Roger.
TOWER: Cancel that.
PLANE: Roger.  Please specify a new destination.
```

A second stage of analysis in the discourse manager deals with implicit references. For example, there are commands such as "remain this frequency" that require contextual knowledge for complete comprehension. Other commands may reference procedures or locations uniquely relevant to a particular airport, or perhaps current weather conditions. Thus, the dispatcher can query the airport/aircraft database for clarification of some commands.

The notions set forth by Grosz *et al* [27] for dealing with the structure of a discourse are particularly relevant in the design of the discourse manager. This preliminary system deals with two dominant intentions: takeoff and landing. These intentional structures determine the response of the system given an input utterance. That is, the utterance is mapped to a corresponding (intermediate) intention, and handlers consult the database to see whether or not all the prerequisites have been met. For example, it does not make sense to issue ground-specific commands to an aircraft in mid-air, and vice versa. Nonsensical or inapplicable commands are rejected and the user is requested to make a new command.

The last stage of the discourse manager is a group of handler functions that actually update the state of the system. In the general case, a very sophisticated (and often intractable) discourse manager is needed to resolve references and to determine the intentional structure of dialogs. However, the air traffic control language has evolved in such a way that much of the ambiguity inherent in human languages has been eliminated. The air traffic control domain is also highly goal-oriented, which permits the assumption that virtually all utterances will be either directive or informative. Without this constraint, the problem of designing a robust discourse manager becomes much more complicated and unpredictable [27] [38] [6] [45] [13].

4.3 Database

In order to interpret and act on commands, the system must refer to information about its own state, as well as the state of the outside world. The system's database acts as a repository for this information. The database contains some pieces of information that are static and do not change throughout the flight, as well as some dynamically changing pieces of information.

An example of static data is most of the information concerning the airport environment. The airport itself is modeled as a set of points on the ground. Each point has a specific (X,Y) coordinate, as well as an adjacency matrix that describes which points are connected to each other via taxiways. Figure 4-5 shows the points and segments that were used in the model of the Laurence G. Hanscom airport in Bedford, Massachusetts [16]. In a system designed to function over a larger geographical area, topographical features and a set of commonly used locational fixes might also be included among the static database features. In an example of the importance of understanding these locational fixes, there is a documented case in which the pilot of an aircraft observed on radar to be flying at an excessively high altitude and in the wrong direction had incorrectly understood a reference to a particular fix, called Maspeth: when instructed to begin a “Maspeth climb,” the pilot instead began a “massive climb” [11]. It was later discovered that the pilot was unfamiliar with the local area, and was not aware of the Maspeth fix. Since the word “massive” is virtually never used in aviation clearances, it is highly unlikely that a pilot (or a speech recognizer) with knowledge of the Maspeth fix would have misinterpreted this clearance.

Dynamically changing database information includes most of the aircraft’s state, as well as non-permanent information about the airport and surrounding environment. The aircraft’s state includes such information as its clearances and queue of commands. Non-permanent airport information might include the current weather conditions, the active runway, and the current landing and takeoff queues.

4.4 Other Features

In order to more accurately simulate appropriate pilot behavior and thus more effectively emulate typical pilot-controller interactions, some non-linguistic features were built into this system. For example, this system contains a variety of ground navigation algorithms. Dijkstra’s Shortest Path algorithm is implemented, but with a slight modification: any ground segment designated as a runway is heavily penalized. Although an airplane will not cross or enter a runway unless it has been given clearance to do so, it is still desirable that the airplanes minimize their time on the runway. This constraint, which mirrors safety procedures practiced by human pilots, reduces the number of runway incursions that are likely to occur. There is also a collision avoidance algorithm, which prevents the controller from having to closely monitor all aircraft traversing the taxiways.

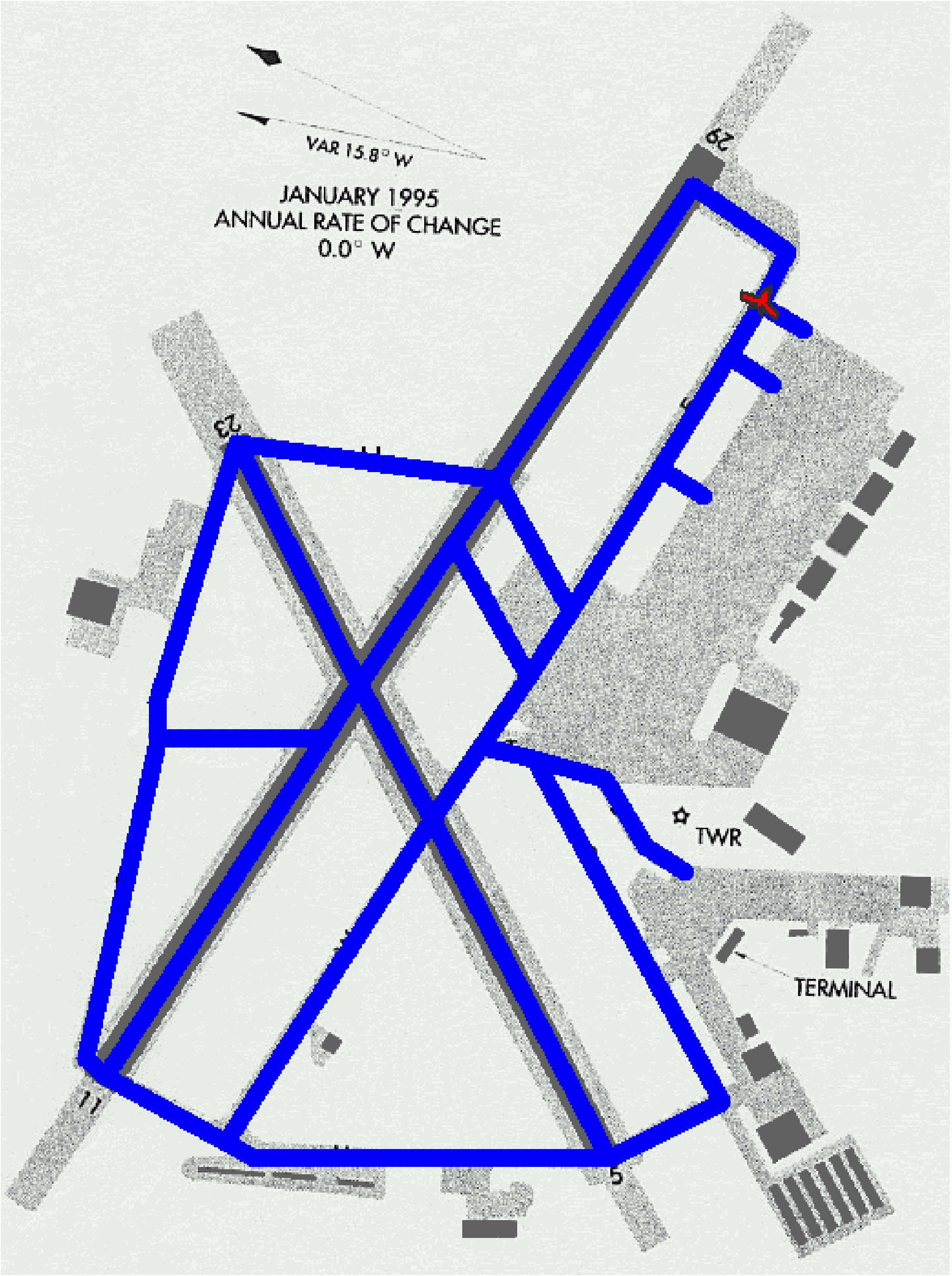


Figure 4-5: Laurence G. Hanscom airfield.

Chapter 5

Demonstration and Comparison With Expanded Corpus

In evaluating the system's performance, it is useful to verify that the system behaves as expected in nominal circumstances, and that it is capable of performing the tasks required of real aircraft. While operation in emergency and non-nominal circumstances it not addressed in this thesis, it is an important topic that should be addressed in future work.

5.1 Example Scenarios

This section illustrates the system's behavior when tested on a number of realistic situations. As the transcripts indicate, both controllability and observability are maintained in all situations.

5.1.1 Scenario 1: Backtracking

The first scenario involves an aircraft taking off. Here, the controller initially intends that the aircraft go to one runway, and subsequently decides to change to another runway. The natural language processing system on the aircraft understands the commands, waits for new orders, and the aircraft proceeds to perform its run-up and announce its readiness for takeoff automatically. Transcript of exchange:

`ALASKA: Ready to taxi.`
`CONTROLLER: Alaska, clear to taxi to runway 5.`
`ALASKA: Roger.` [Begins to taxi]
`CONTROLLER: Cancel that.`
`ALASKA: Roger. Awating a new destination.`
`CONTROLLER: Taxi to twoniner via echo.`
`ALASKA: Roger.` [Taxis to runway 29 via echo]
`ALASKA: Runup completed. Ready for departure.`
`CONTROLLER: Alaska, hold short.`
`ALASKA: Roger.`

`CONTROLLER: Alaska, position and hold.`
`ALASKA: Position and hold.` [Taxis to position]
`    Ready for departure.`
`CONTROLLER: Clear for takeoff.`
`ALASKA: Roger.` [Takes off]

5.1.2 Scenario 2: Controller-enforced Queuing

The second scenario involves two aircraft (Horizon and United) who both want to land. The controller gives precedence to the United over the Horizon, and thus Horizon lands only after the United has landed. Transcript of exchange:

`HORIZON: Inbound`
`UNITED: Inbound`
`CONTROLLER: Horizon, you're number two for landing behind a United.`
`United, you're clear for landing runway 29.`
`UNITED: Roger. Preparing to land on runway 29.` [Horizon does not prepare.]
`CONTROLLER: Horizon, clear to land runway 29.`
`HORIZON: Roger. Preparing to land on runway 29.`

5.1.3 Scenario 3: Automatic Queuing

The third scenario involves two aircraft (Alaska and Boeing) who both need to cross the same intersection. The controller had previously instructed Alaska to hold for Boeing, thus giving priority to the Boeing in case of conflict. Transcript of exchange:

`CONTROLLER: Alaska, hold for the Boeing.`
`ALASKA: Roger.`

If the two aircraft reach the intersection at about the same time, Alaska will wait until Boeing passes.

5.1.4 Scenario 4: Controller Error

In the fourth scenario, Alaska is heading toward runway 29 when the controller specifies a particular taxiway. However, the controller says the wrong runway number (or, the speech recognizer heard the wrong number). The discourse manager catches that, and the aircraft asks for clarification. Once the controller issues the new commands, the aircraft is able to perform its run-up and take off. Transcript of exchange:

`ALASKA: Ready to taxi.`
`CONTROLLER: Alaska, clear to runway 29.`
`ALASKA: Roger.` [Begins to taxi to 29]
`CONTROLLER: Taxi to runway 10 via echo.`
`ALASKA: I do not think there is such a runway.`

```
CONTROLLER: Taxi to runway 29 via echo.
ALASKA: Roger.
```

5.2 Comparison with Expanded Corpus

The real test of a language processing system comes when it is exposed to realistic material on which it was not trained. It is therefore useful to compare the range of inputs accepted by this system with the actual commands passed from controllers to pilots in practice. For this comparison, a corpus of air traffic control exchanges from the Linguistic Data Consortium [23] was examined.

Figures 5-1 – 5-7 show the relative frequency of commands given by several types of controllers over twenty-minute intervals at Washington National Airport and Dallas-Fort Worth Airport. There is a predominance of primitive commands such as vectoring, all which are handled by this system. Most composite commands are also handled by this system, although the particular issues of following traffic and utilizing navigational aids are not addressed. Many advisory statements given to human pilots do not have immediate impact on flight, and may not be appropriate for unmanned aircraft. However, these advisory statements may be useful in interpretation of "party line information," which will be discussed further in Chapter 6. Finally, almost all queries and qualifiers can be handled by this system, although those pertaining to visual traffic separation and inappropriate for non-vision-based systems, and requests to expedite a procedure are not currently handled.

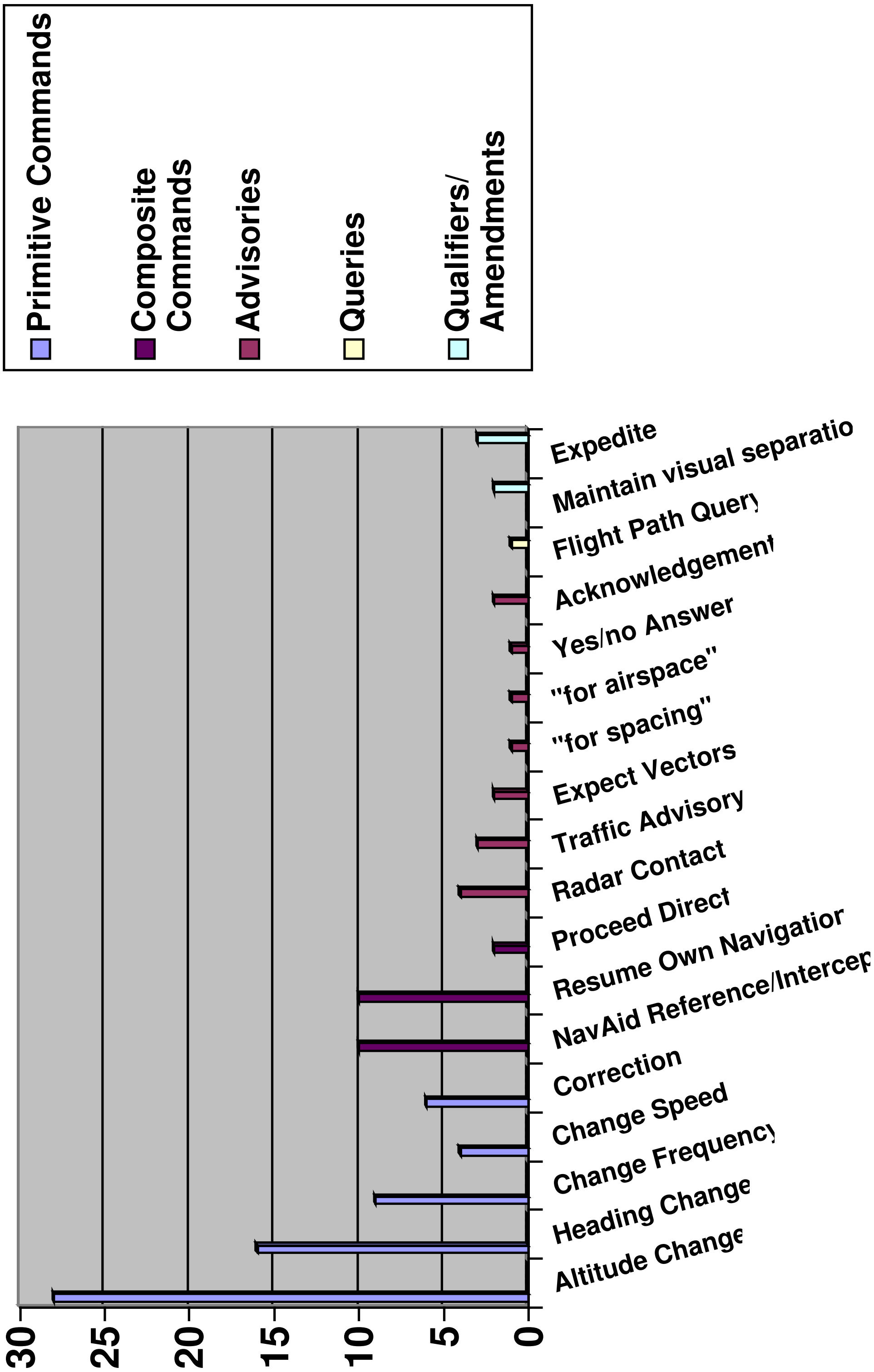


Figure 5-1: Types of commands given by a departure controller at Washington National Airport during normal operation.

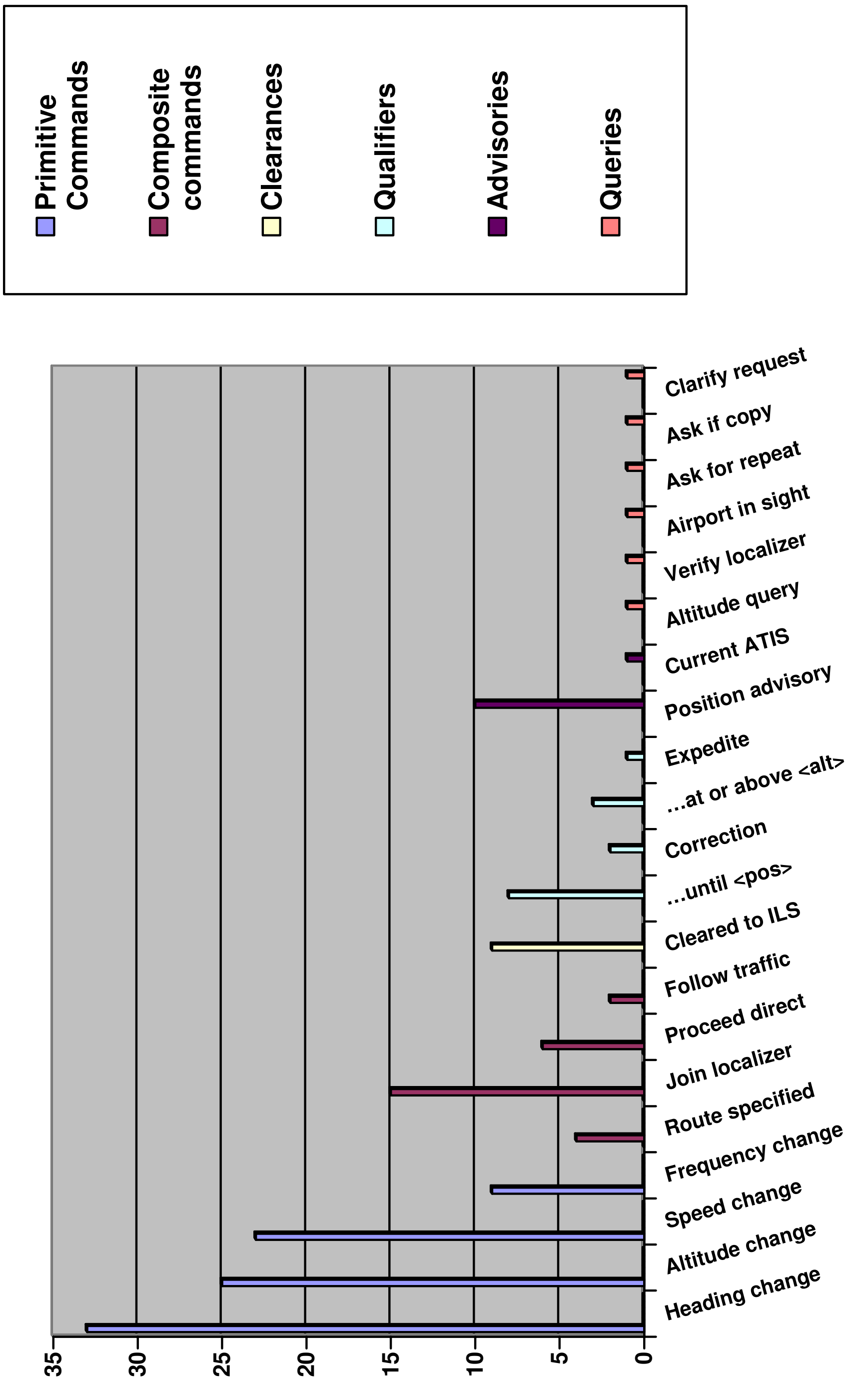


Figure 5-2: Types of commands given by a feeder controller at Washington National Airport during normal operation. A feeder controller is responsible for handling aircraft within about 50 miles of the airport, until they are handed to a local controller.

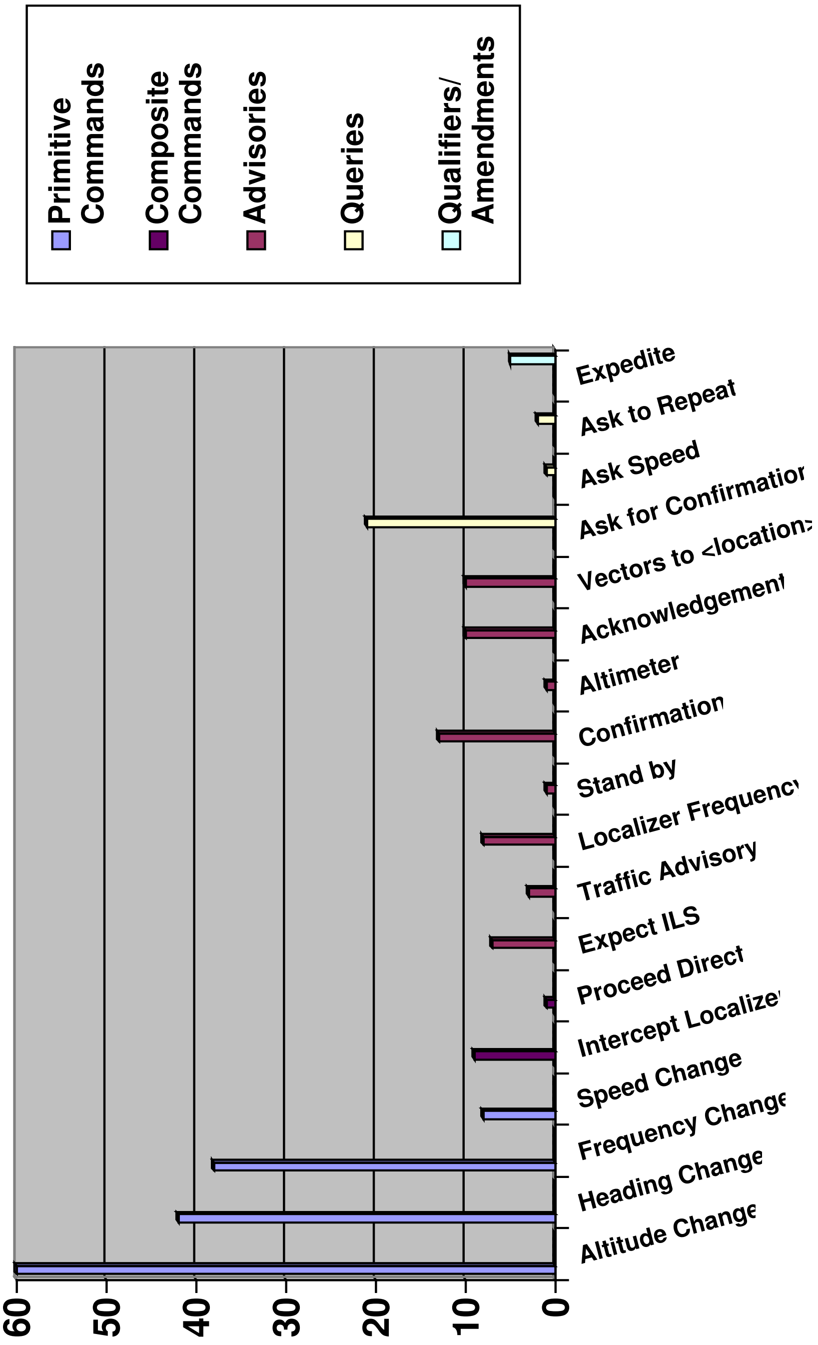


Figure 5-3: Types of commands given by a feeder controller at Dallas-Fort Worth Airport during normal operation.

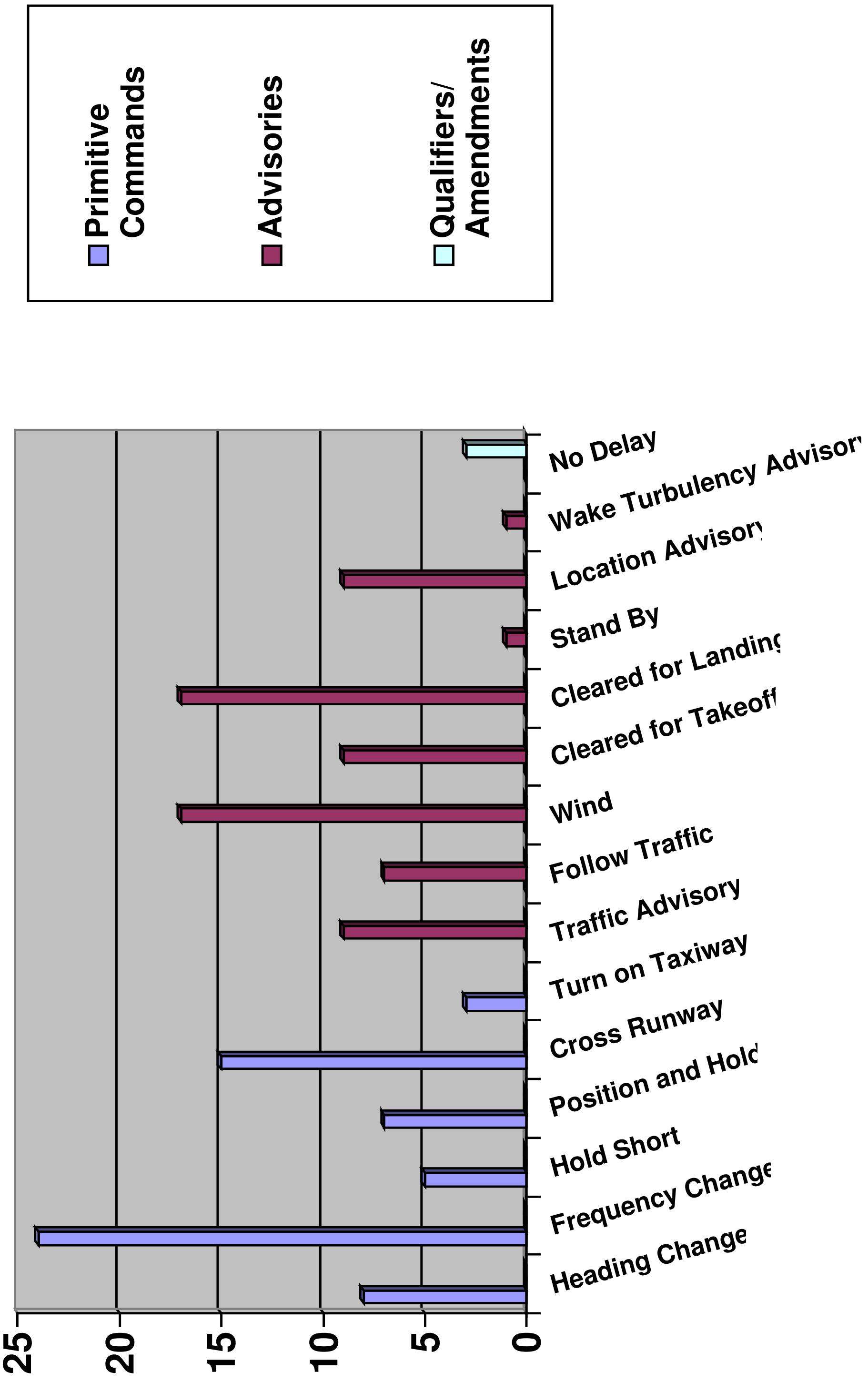


Figure 5-4: Types of commands given by a local controller at Dallas-Fort Worth Airport during normal operation.

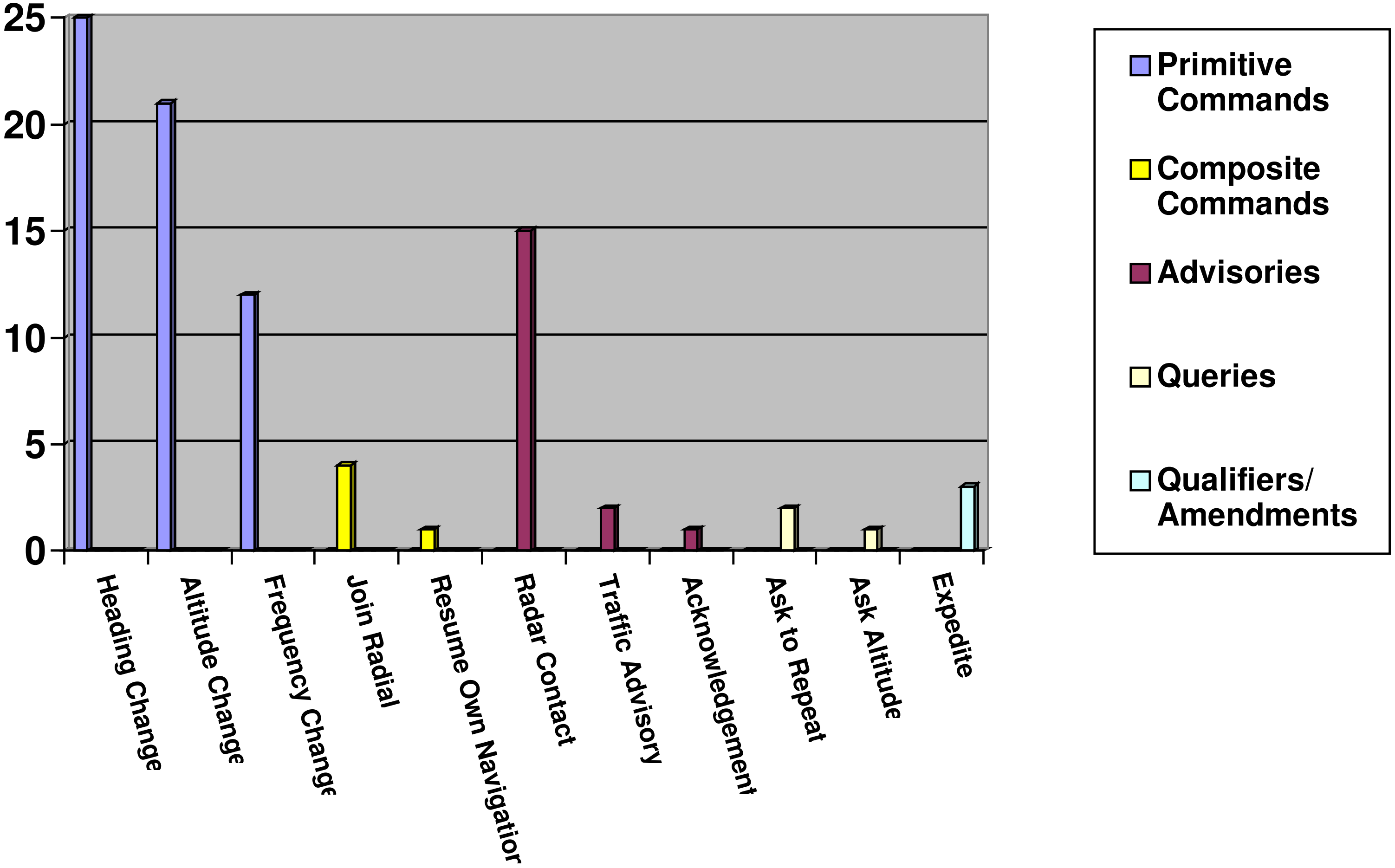


Figure 5-5: Types of commands given by a departure controller at Dallas-Fort Worth Airport during normal operation.

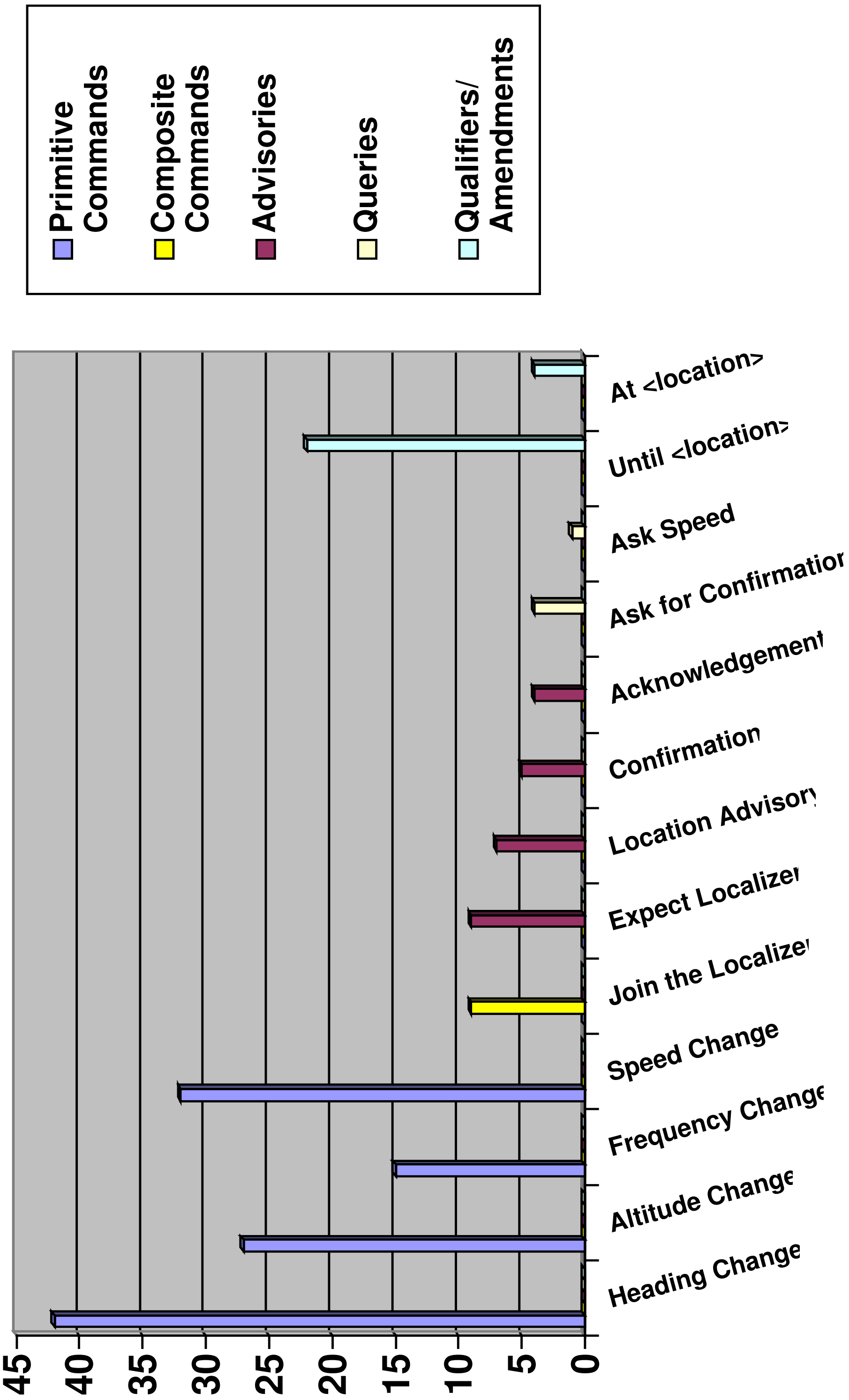


Figure 5-6: Types of commands given by an arrival controller at Dallas-Fort Worth Airport during normal operation.

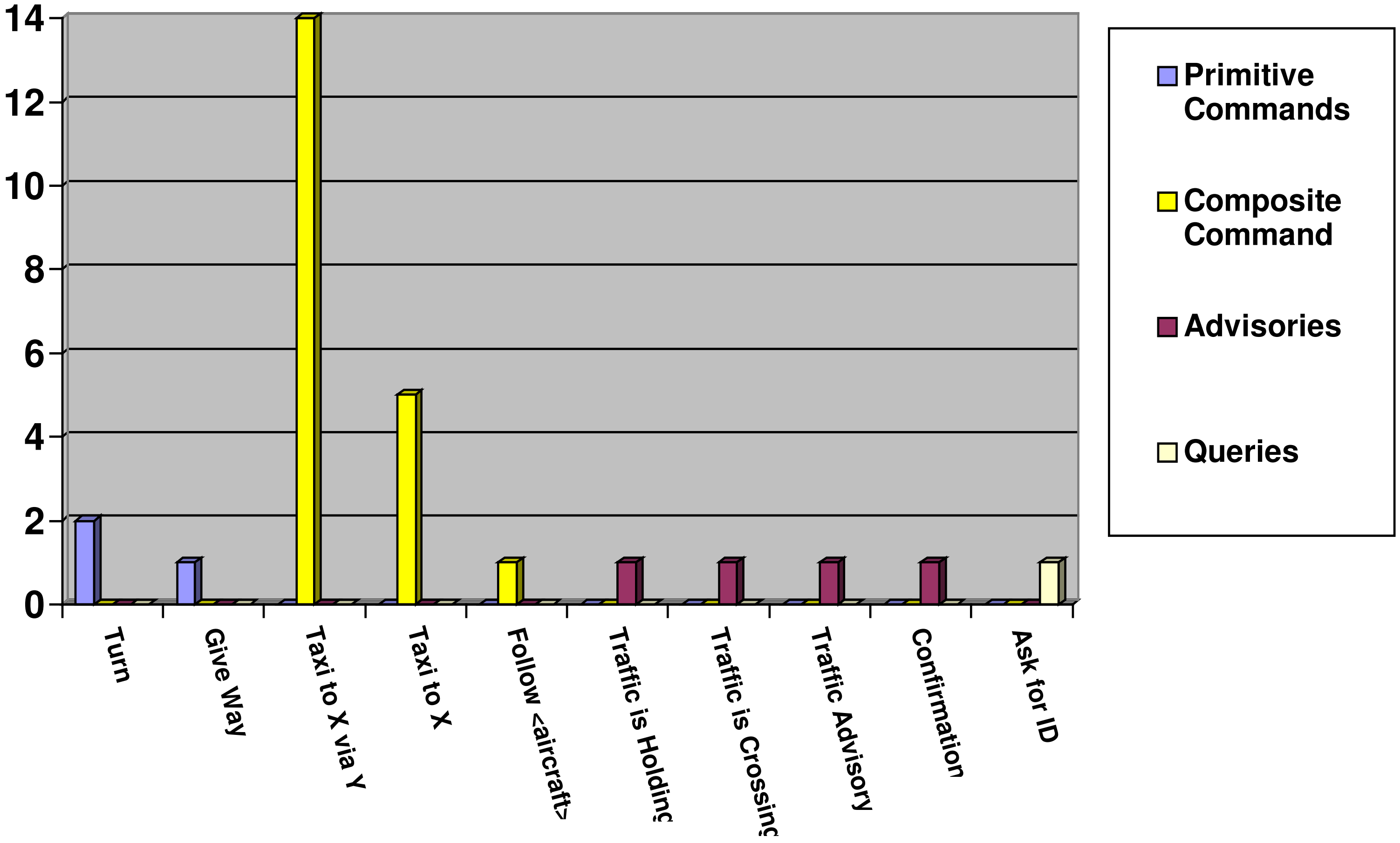


Figure 5-7: Types of commands given by a ground controller at Dallas-Fort Worth Airport during normal operation.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis has outlined some of the challenges and opportunities present in the application of natural language processing technology to the guidance of unmanned aerial vehicles. It has also described a basic natural language interface that, despite its simplicity, nevertheless captures much of the behavior of a real system. Thus, it would seem that with the proper extensions and improvements to this work, natural language processing may be a promising technology for improved UAV guidance interfaces.

Many issues remain to be investigated before a system of this nature may be successfully deployed. However, if the potential benefits of the natural language processing technology described in this thesis are realized, the utility of UAVs will be greatly increased. Aside from allowing current users of UAVs to interact more easily with the vehicles, this technology could open up entire new venues for the use of UAVs. Examples of potential areas for expanded use of UAVs include scientific data collection missions, search and rescue operations, military reconnaissance and surveillance, urban operations, film making, agricultural endeavors, and many others.

It should also be noted that the applications of this type of technology are not limited to unmanned aircraft. Improved dialog systems would benefit a number of other applications involving semi-autonomous and knowledge-based systems. Furthermore, improved capability in the field of language processing carries ramifications extending beyond the pragmatic engineering issues and into some fundamental ideas about human cognition and machine intelligence. Language acquisition is regarded as one of the most fundamental features of human intelligence, paving the way for more advanced capabilities in reasoning and deduction. The ability to reason and use language in a humanlike fashion has long been seen as a hallmark of artificial intelligence, made concrete by Alan Turing's famous criterion for defining a truly "intelligent" artificial agent. In an age where computers are becoming constant features in everyday life, it is imperative that these devices be capable of integrating seamlessly into the fabric of human interaction, and this is at least as true in the increasingly computerized and automated world of aviation as it is in every other area of life.

6.2 Future Work

Like much of the research that takes place in the early stages of a concept's development, this thesis perhaps raises more questions than it answers. Besides improvements that might be made to the particular system described in this thesis, there are also a number of supporting and complementary technologies that may be developed, and several issues that need to be addressed before a hardware implementation of such a system may be realized.

6.2.1 Interface Improvements

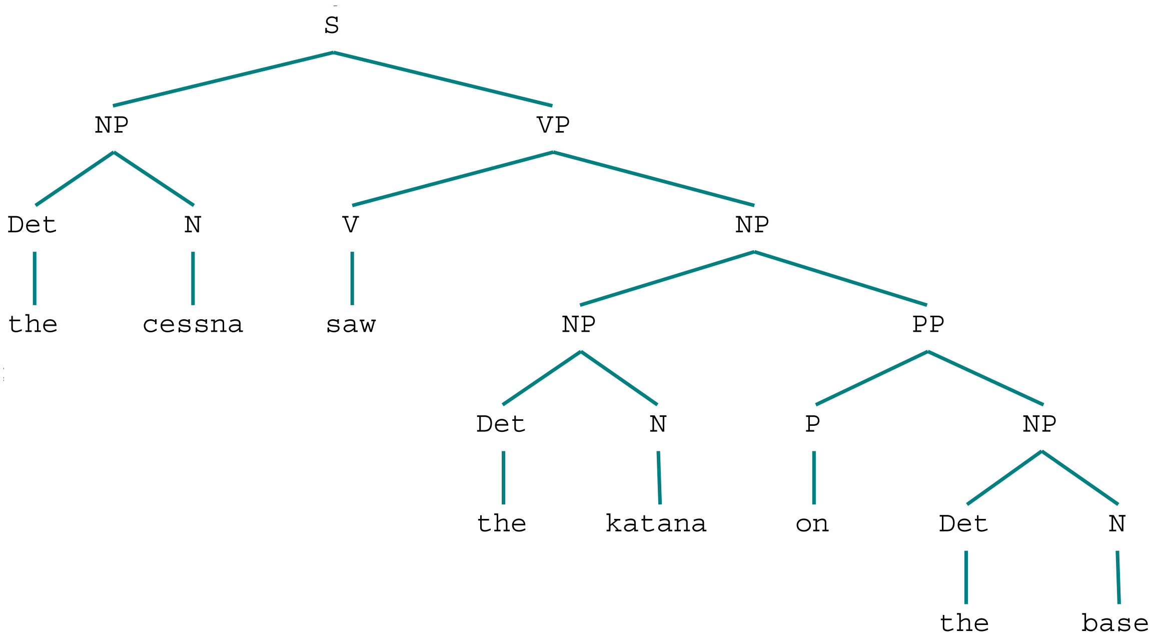
Future additions and improvements to this system might include a multiple-aircraft interfacing capability that allows the controller to address all or part of a fleet of aircraft, as well as an individual craft; a database manager capable of incorporating information extracted from such sources as automated weather and airport advisories and transmissions between other aircraft ("party line information") as well as transmissions to that aircraft; and of course further additions to the corpus of known sentences and a more sophisticated discourse manager capable of time-sensitive discourse and improved and expanded intentional inference.

The system described in this thesis also makes no effort to resolve ambiguity. This did not have a significant operational impact on the system because most ambiguity that may exist in a dialog occurs during third-person reference, not direct imperative communication. However, the incorporation of party-line information into the system's logical analyzer would necessitate disambiguation. For example, a sentence such as that shown in Figure 6-1 would require contextual information to decipher properly.

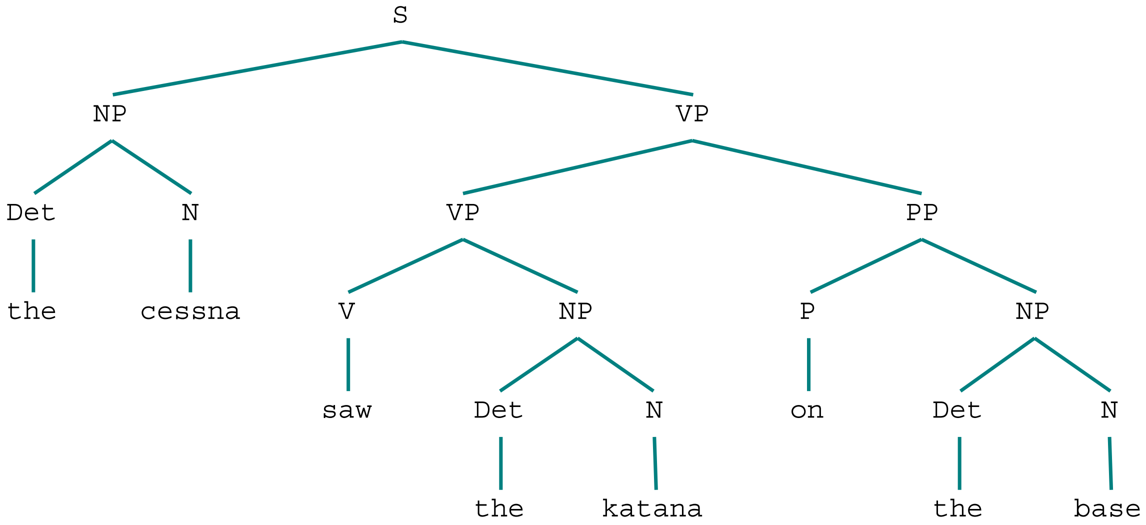
6.2.2 Supporting and Complementary Technologies

Some existing or currently-developing technologies would be very effective when paired with the ideas put forth in this thesis. For example, the integration of optimal path planning algorithms for aircraft in flight and the development of cooperative strategies for multiple-aircraft operations could allow a single operator to easily and efficiently control many aircraft.

There is also complementary theoretical research that could build on this work. An area for further research about the structure of the air traffic control language would involve the notion of memory in air traffic control dialogs. How long should information be retained in the course of a dialog, and when should it be considered obsolete? To answer this question, transcripts of approximately 70 hours of recorded air traffic control exchanges obtained from the Linguistic Data Consortium air traffic control corpus [23] will be examined for references to previous commands and advisories. For validation, these exchanges may be augmented with information such as that which may be obtained from the Aircraft Communication Addressing and Reporting System (ACARS). The ACARS maintains a record of aircraft operations and may contain flight information for some of the same aircraft whose communications



(a) The semantic interpretation of this parse is that the Cessna saw the Katana when the Katana was on the base (the Cessna's location is unspecified).



(b) This parse indicates that the Cessna was on the base when it saw the Katana (the Katana's location is unspecified).

Figure 6-1: Two valid parsings of an ambiguous sentence, "The Cessna saw the Katana on the base."

are recorded in the LDC transcripts {citeacars. The author expects to see patterns based on time difference between utterances, situational context of utterances, and perhaps relative importance or urgency of utterances.

There is also an important human factors problem to be addressed – how effective is voice control of UAVs? The author suspects that a natural language interface would have several advantages over more conventional means of UAV control, especially in situations in which a single controller is dealing with both manned and unmanned aircraft. However, this advantage should be quantified through experiments with various control schemes and in a variety of situations in a simulated environment. Control systems should be evaluated with respect to the factors outline in Chapter 2, as well as such factors as efficiency, safety, ease of use by air traffic controllers, and the ease with which they might be incorporated into existing air traffic control systems.

Another area for further research is in dialog-based aircraft state estimation. For an aircraft operating in a shared airspace, many external factors dictate the ways in which that aircraft should operate and communicate. For instance, a verbose status report should not be given to an air traffic controller who is already trying to deal with many other aircraft at once. The following quotation illustrates this concept in a rather extreme situation; note the discretion with which Virgin 12 speaks to the controller when an emergency situation requiring the controller's full attention is detected on Swissair 111:

> SWISSAIR 111: Swissair one-eleven heavy is declaring Pan Pan Pan. We have, uh, smoke in the cockpit. Uh, request immediate return, uh, to a convenient place, I guess, uh, Boston.
> MONCTON CONTROLLER: Swissair one-eleven, roger . . . turn right proceed . . . uh . . .you say to Boston you want to go.
> SWISSAIR 111: I guess Boston . . . we need first the weather so, uh, we start a right turn here. Swissair one-one-one heavy.
> MONCTON CONTROLLER: Swissair one-eleven, roger, and a descent to flight level three-one-zero. Is that OK?
> SWISSAIR 111: Three-one-zero. (Unintelligible words obscured by a noise. Possibly the noise associated with donning oxygen masks.) Three-one-zero . . . one-one heavy.
> MONCTON CONTROLLER: Swissair one-eleven, Centre.
> SWISSAIR 111: Swissair one-eleven heavy, go ahead.
> MONCTON CONTROLLER: Uh, would you prefer to go into Halifax?
> SWISSAIR 111: Uh, standby.
> VIRGIN 12: **Moncton, Virgin twelve will be standing by.**
> MONCTON CONTROLLER: **Virgin twelve, roger, standby.**
> SWISSAIR 111: Affirmative for Swissair one-eleven heavy. We prefer Halifax from our position.
> MONCTON CONTROLLER: Swissair one-eleven, roger. Proceed direct to Halifax.

Similarly, an aircraft should proceed cautiously if an emergency situation is detected at or near an airport. Ideally, these operational protocols and nuances of etiquette should be observed by unmanned aircraft as readily as they are by human pilots. For this to occur, the UAV will need to have an accurate model of the other aircraft that are operating near it, including their approximate positions and operational states. Because in many cases the only contact that aircraft have with one another is through communication channels, it is reasonable to attempt to make state estimations based on dialogs between aircraft and air traffic controllers.

Research in this area will attempt to identify the most effective techniques for dialog-based state estimation in the domain of air traffic control. Knowledge of air traffic control procedures and phraseology will be used in conjunction with existing modeling, labeling, and classification algorithms in order to distill the features of the air traffic control language that are relevant to state estimation, as well as the best means for processing these features and making an estimation. Specifically, hidden Markov modeling techniques will be used with a variety of features of ATC utterances, such as command given, speaker and receiver of utterance, and time between utterances, in order to identify those features that play a useful role in state estimation. Other classification algorithms may be evaluated in a similar manner, and kinematic models with propagated uncertainty may be incorporated.

The area of research to which this idea bears the most resemblance is that of discourse analysis, specifically dialog act tagging. In such research, utterances are analyzed across a set of features, and an appropriate labeling scheme for these utterances is attempted. The most notable research in this area with respect to spoken dialog was undertaken by Stolcke *et al* [61], who set out to model conversational dialog as a series of "speech acts" such as Statement, Question, Yes Answer, etc. They used such features as lexical and prosodic cues, proximity, and overall discourse structure as parameters in a hidden Markov model with the discourse model as states and the dialog acts as observations. While there are some differences between this research and the proposed project (some differences having to do with feature selection, tag set selection, and the task-oriented nature of air traffic control, and others having to do with the amount of data used - a full conversation vs. a running total as the dialog proceeds), there are also many similarities and opportunities for elaboration, especially where HMM training techniques are concerned.

Some related research examines dialog acts in task-oriented dialogs that more closely mirror air traffic control exchanges than do the Switchboard corpora processed by Stolcke *et al.* For example, Wright [73] performed analysis on a dialog between two individuals attempting to reach a common goal of understanding a map, while Reithinger and Klesen [54] analyzed dialogs between two individuals trying to solve a scheduling problem. Both of these domains are quite similar to air traffic control in that they are task-oriented, and virtually all exchanges take place between two individuals. Furthermore, the techniques used in these studies do not depend on prosodic information, so it is feasible to implement some of the same strategies in a transcript-oriented project. Finally, work done by Barzilay and Lee [1] is pertinent, since it attempts to classify prose consisting of many sentences rather than one – a notion that may well prove useful in air traffic control dialogs, since an aircraft state will typically (though not always) remain unchanged through several utterances.

A well-established fixture in biological research [14], Hidden Markov Modeling techniques are also ubiquitous in language processing. In fact, HMMs are essential in the dialog act research described above, as well as in related areas such as part of speech tagging and disambiguation. The prospect of using HMMs in the domain of air traffic control is particularly attractive due to the underlying assumption that the probability of a state in a Markov chain can be determined based on some finite (and usually small) number n previous states, a notion that is intuitively likely to be true of a terse and formal language such as air traffic control.

Other types of classifiers also exist in language processing. These classifiers typically utilize relatively shallow discourse markers, such as cue words, as features. The air traffic control language may be a good candidate for such techniques. Types of classifiers include neural networks [35], decision trees [37], and others.

In order to incorporate the continuous and real-time aspects of flight, a kinematic model may be combined with the language processing techniques described above in order to form a hybrid discrete-continuous model. Such utterances as position reports and primitive commands could be used to "fix" an aircraft at a particular point and establish an estimated trajectory for it, while uncertainty about this trajectory may be propagated over time to produce a region in which an aircraft may be located. This region may then be updated with subsequent commands.

6.2.3 Implementation Issues

In addition to the voice recognition required for a hardware implementation, there are also other practical considerations to take into account. For example, a UAV utilizing a language processing system will need to be able to decide when to start and stop listening for a command. At first glance the standard air traffic control procedure appears adequate: When the name of the UAV is heard, this marks the beginning of a command utterance. A long pause may indicate the end. However, this strategy is not foolproof - for example, it may happen that the UAV hears someone talking *about* it rather than *to* it. Thus, a more robust algorithm, or possibly a slight modification to air traffic control procedures for UAVs, is required.

There is also another layer of ambiguity present in spoken language that is not present in written language: ambiguity based on word sound similarities. For example, the following is a documented case of communication ambiguity [11]:

> CONTROLLER: *clears aircraft to descend* "two four zero zero"
> PILOT: *heard a clearance to descend "to four zero zero;" reads back clearance as* "OK. Four zero zero."

Having understood the controller's instruction as a directive to descend to 400 feet (not 2400 feet, as the controller intended), the pilot proceeded to do so. It should be noted that there were multiple opportunities for this particular ambiguity to be resolved. First, the pilot might have noticed the word-sound ambiguity and asked the controller for clarification. Secondly, the pilot might have found a command to descend to 400 feet anomalous, since the minimum legal altitude for flying an aircraft is 500 feet above ground level in unpopulated areas, and 1000 feet above ground level in populated areas. Flight below 500 feet generally takes place only during takeoff and landing, and a landing descent is usually done at the pilot's discretion rather than on the controller's command. Given that the aircraft was apparently at an altitude of more than 2400 feet prior to the command, it is unlikely that it was in an appropriate position to begin this rather drastic descent for landing. Thus, it is reasonable to expect that the pilot might question such an unusual command, and might voice his or her concerns to the controller. Finally, the pilot's read back of the incorrect altitude provided an opportunity for the controller to make a correction. Despite these many opportunities for the pilot's misunderstanding to be caught, it sill resulted in a fatal accident. Thus, if there are two possible speech taggings for a given utterance, and both are semantically reasonable, the system should be capable of generating a response to the user asking for clarification of the ambiguity.

Another area for further development is in command boundary detection. While the implementation described in this thesis takes one command at a time, in reality commands appear in rapid sequence, such as these taken from air traffic control transcripts:

> ‘‘CHEYENNE NINER THREE XRAY IS EIGHT MILES FROM OXONN TURN LEFT HEADING ZERO FOUR ZERO INTERCEPT THE LOCALIZER AT TWO THOUSAND FIVE HUNDRED FEET CLEARED I L S RUNWAY THREE SIX APPROACH’’

> ‘‘AMERICAN TWO THIRTY SIX DESCEND AND MAINTAIN FOUR THOUSAND WHEN YOU GET TO FOUR THOUSAND REDUCE TO TWO HUNDRED AND TEN KNOTS’’

> ‘‘THREE THREE NOVEMBER JULIET IS TEN FROM OX CORRECTION ONE FIVE MILES FROM OXONN INTERCEPT THE LOCALIZER AT OR ABOVE TWO THOUSAND FIVE HUNDRED CLEARED I L S RUNWAY THREE SIX APPROACH’’

This boundary detection may not be possible to do using only parsing; for example, the sentence "The Cessna heard the Katana contact the tower" (i.e., Cessna confirms that Katana's radio is working) is syntactically identical to "The Cessna saw the Katana; contact the tower," which is a traffic advisory followed by a command to contact the tower.

Finally, while a simulation environment is useful for testing the system's behavior in normal conditions, it is not appropriate for examining behavior in emergency or unusual situations. Thus, emergency procedures will need to be considered and developed in tandem with hardware development and implementation.

Bibliography

[1] Barzilay, Regina; Lee, Lillian. "Catching the drift: Probabilistic content models, with applications to generation and summarization." *Proceedings of HLT/NAACL 2004*, pp. 113–120. 2004

[2] Bertsekas, Dimitri P. *Dynamic Programming and Optimal Control, Second Edition.* Athena Scientific, 2000.

[3] Blazakis, Jason. "CRS Report for Congress: Border Security and Unmanned Aerial Vehicles." January 2004.

[4] Bone, Elizabeth; Bolkcom, Christopher. "Unmanned Aerial Vehicles: Background and Issues for Congress." 2003.

[5] Bush, George W. "President Speaks on War Effort to Citadel Cadets." Remarks by the President. December 2001.

[6] Carberry, S. *Plan Recognition in Natural Language Dialog.* MIT press, 1990.

[7] Chandioux, J. "METEO: un systeme operationnel pour la traduction automatique des bulletins meteorologiques destines au grand public." *Meta*, 21, pages 127-133. 1976.

[8] Chomsky, N. "Three models for the description of language." *IRE Transactions on Information Theory*, 2 (3), 113-124. 1956.

[9] Chomsky, N. "On Certain Formal Properties of Grammars." *Information and Control, 2*, 136-137. 1959.

[10] Churcher, Gavin E.; Ateall, Eric S. and Souter, Clive, "Dialogues in Air Traffic Control," *Proceedings of the 11th Twente Workshop on Language Technology*, 1996.

[11] Cushing, Stephen. *Fatal Words: Communication Clashes and Aircraft Crashes.* The University of Chicago Press, 1994.

[12] Department of Defense. "Program Acquisiton Costs By Weapon System". 2003.

[13] Doherty, Patrick; Granlund, Gosta; Kuchcinski, Krzystof; Sandewall, Erik; Nordberg, Klas; Skarman, Erik and Wiklund, Johan, "The WITAS Unmanned Aerial Vehicle Project", *Proceedings ECAI*, 2000.

[14] Durbin, Richard; Eddy, Sean; Krogh, Anders; Mitchison,Graeme. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids.* Cambridge University Press, 1998.

[15] Earley, J. "An Efficient Context-Free Parsing Algorithm." *Communications of the ACM*, 6(8), pages 451-455, 1970.

[16] Federal Aviation Administration. *Airport/Facility Directory, Northeastern U.S.*, National Aeronautical Charting Office, U.S. Department of Transportation, 2002.

[17] Federal Aviation Administration. "User Benefits of Two-Way Data Link ATC Communications: Aircraft Delay and Flight Efficiency in Congested En Route Airspace." DOT/FAA/CT-95/4. Data Link Benefits Study Team Report. U.S. Department of Transportation, Washington, DC. 1995.

[18] Federal Aviation Administration. "Data Link Benefits Study Team Report." 1996.

[19] Federal Aviation Administration. *2002 Federal Aviation Regulations and Aeronautical Information Manual.* U.S. Department of Transportation, 2002/

[20] Freudenrich, Craig C. "How Air Traffic Control Works." *http://www.howstuffworks.com*, 2004.

[21] Gates, William. "Remarks at the Gartner Symposium." 1997.

[22] Gent, R.N. and Van, H.W. "Human Factors Issues with Airborne Data Link." NLR Technical Publication 95666L. National Aeronautics Laboratory, Amsterdam, Netherlands, 1996.

[23] Godfrey, John J. *Air Traffic Control Complete Corpus.* LDC94S14A. CD- ROM. Philadelphia: Linguistic Data Consortium, 1997.

[24] Gorin, A.L.; Parker, B.A.; Sachs, R.M.; and Wilpon, J.G. "How May I Help You?" AT&T Research, 1997.

[25] Guardino, John R. "Manning and Unmanning Army Aviation." *Rotor and Wing.* December, 2002.

[26] Grosz, B., "The Representation and Use of Focus in Dialogue Understanding." Ph.D. thesis, University of California, Berkeley, 1977.

[27] Grosz, Barbara and Sidner, Candance L., "Attention, Intentions, and the Structure of Discourse", *Computational Linguistics*, 12(3), pages 175-204, 1986.

[28] Heppenheimer, T. A. *Countdown: A History of Space Flight.*John Wiley and Sons, 1999.

[29] Hobbs, J.R. "Coherence and Coreference", *Cognitive Science*, 3, pages 67-90. 1979

[30] Huybregts, R. *The Weak Inadequacy of Context-Free Phrase Structure Grammars.* In de Haan, G.; Trommele, M.; and Zonneveld, W. (Eds.), *Van Periferie naar Kern.* Foris, Dordrecht, 1984.

[31] Jacobsen, D. M. "Unmanned Aerial Vehicles – the Key to Effective Situational Awareness in Littoral Operations." 2002.

[32] Jurafsky, D. and Martin, J.H., *Speech and Language Processing*, Prentice Hall, 2000.

[33] Richard Kebabjian. *http://www.planecrashinfo.com.* 2004.

[34] Kerns, K. "Human Factors in the ATC/Flight Deck Integration: Implications of Data Link Simulation Research." MP 94W0000098. MITRE Corporation, McLean, VA, 1994.

[35] Kipp, M. "The Neural Pathway to Dialogue Acts." *Proceedings of the 13th ECAI*, 1998.

[36] Lemon, Oliver; Bracy, Anne; Gruenstein, Alexander; Peters, Stanley. "The WITAS Multi-Modal Dialogue System I." In proceesings from *Eurospeech*, 2001.

[37] Litman, D. "Cue Phrase Classification Using Machine Learning." *Journal of Artificial Intelligence Research*, 1996.

[38] Litman, D.J. and Allen, J.F. "A Plan Recognition Model for Subdialogues in Conversation." *Cognitive Science*, 11, pages 163-200, 1987.

[39] Mayer, Daryl. *AWACS Voice Recognition May Enhance Accuracy.* AFMC News Service Release 0334. March 22, 2004.

[40] National Geographic Explorer, CNBC. *The Pigeon Murders.* Documentary. 2000.

[41] Newcome, Laurence. *A Brief History of Unmanned Aviation.* American Institute of Aeronautics and Astronautics, 2004.

[42] Nolan, Michael S. *Fundamentals of Air Traffic Control*, Third Edition. Brooks/Cole Publishing Company, 1999.

[43] Office of the Secretary of Defense. "UAV Roadmap 2002-2027". December 2002.

[44] Office of the Secretary of Defense. “Unmanned Aerial Vehicles Roadmap 2000-2025.” 2001.

[45] Passonneau, R. and Litman, D.J., “Intention-based Segmentation: Human Reliability and Correlation with Linguistic Cues”, ACL 93, Columbus, Ohio, 1993.

[46] Pike, John. “RAH-66 Comanche.”
http://www.globalsecurity.org/military/systems/aircraft/rah-66.htm

[47] Pilot interviews, Training Squadron VT-86. Personal communication, 2004.

[48] *Private Pilot Manual*, Jeppenson Sanderson, Inc., Englewood, CO, 80112-5498, 2001.

[49] Programme for Harmonised Air Traffic Management Research in Eurocontrol PD1, “Evaluation of a Label Oriented HMI for Tactical Data Link Communication in ATC.” Document 96-70-25, Brussels: Eurocontrol, 1996.

[50] Pullum, G. K. and Gazdar, G. “Natural Languages and Context-Free Languages.” *Linguistics and Philosophy, 4*, 471-504. 1982.

[51] Quigley, Morgan; Goodrich, Michael A.; Beard, Randal W. “Semi-Autonomous Human-UAV Interfaces for Fixed-Wing Mini-UAVs.”

[52] Ramsey, James W. “UAVs: Out of Uniform.” Avionics Magazine, March 2004.

[53] Reeves, B. and Nass, C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places.* Cambridge University Press, Cambridge, 1996.

[54] Reithinger, Norbert; Klesen, Martin. “Dialog Act Classification Using Language Models.” *Proceedings of the 5th European Conference on Speech and Communication Technology*, 1997.

[55] Schoenung, Susan. “White Paper: UAV Over-the-Horizon Disaster Management Projects.” Nasa Ames Research Center, 2000.

[56] Schouwenaars, T.; Feron, E.;How, J. “Hybrid Architecture for Receding Horizon Guidance of Agile Maneuvering Autonomous Rotorcraft,” *16th IFAC Symposium on Automatic Control in Aerospace*, June 2004.

[57] Schouwenaars, T.; How, J; Feron E. "Receding Horizon Path Planning with Implicit Safety Guarantees," *American Control Conference*, June 2004.

[58] Shieber, S. M. "Evidence Against the Context-Freeness of Natural Language." *Linguistics and Philosophy, 8*, 333-343, 1985.

[59] Sia, Richard. "Pilotless Aircraft Makers Seek Role for Domestic Uses." National Journal's Congress Daily, December 2002.

[60] Space Daily. "NASA Uses Remotely Piloted Airplane To Monitor Grapes." August 20, 2003.

[61] Stolcke, Andreas; Ries, Klaus; Coccaro, Noah; Shriberg, Elizabeth; Bates, Rebecca; Jurafsky, Daniel; Taylor, Paul; Martin, Rachel; Van Ess-Dykema, Carol; Meteer, Marie. "Dialog Act Modeling for Automatic Tagging and Recognition of Conversational Speech." *Association for Computational Linguistics*, 2000.

[62] Talotta, N.J. et al. "Operational Evaluation of Initial Data Link En Route Services, Volume 1." Report No. DOT/FAA/CT-90/1, I. Federal Aviation Administration, U.S. Department of Transportation, Washington, D.C. 1990.

[63] Talotta, N. J. et al. "Controller Evaluation of Initial Data Link Terminal Air Traffic Control Services: Mini-Study 2, Volume I." Report No. DOT/FAA/CT-92/2, I. Federal Aviation Administration, U.S. Department of Transportation, Washington, D.C. 1992.

[64] Talotta, N. J. et al. "Controller Evaluation of Initial Data Link Terminal Air Traffic Control Services: Mini-Study 3, Volume I." Report No. DOT/FAA/CT-92/18, I. Federal Aviation Administration, U.S. Department of Transportation, Washington, D.C. 1992.

[65] United States Air Force Scientific Advisory Board; Ray Johnson, Malcolm ONeill, Peter Worch, Greg Zacharias, Brian Hunt, VADM (ret) Lyle Bien, Jim Lang, William Lawler, Steve May, Chris Mitchell, Ken Pedersen, Heidi Shyu, Phil Soucy, Mike Yarymovych, Stephen Cross, Mica Endsley, Jeffery Erickson, Matthew Ganz, Teresa Lunt, Robin Murphy, Shankar Sastry, Michael Shatz, Ed Brady, John Entzminger, Brian Argrow, Harry Berman, Robert Byer, Claude Canizares, Armand Chaput, Tom Cruse, Gary Denman, Hamish Fraser, Wally Hoff, Don Kenney, Robert MacCormack, Richard Murray. "Unmanned Aerial Vehicles in Perspective: Effects, Capabilities, and Technologies". July 2003.

[66] United States Air Force Scientific Advisory Board. *New World Vistas: Air and Space Power for the 21st Century.* 1995.

[67] United States Department of Defense. *Unmanned Aerial Vehicle Master Plan.* 1991.

[68] United States Navy, Naval Air Training Command, *Voice Communications: Student Guide.* CNATRA P-806 (REV. 4-98) PAT. NAS Corpus Christi, TX, 1998.

[69] Van de Vegte, John. *Feedback Control Systems*, Third Edition. Prentice Hall, 1994.

[70] Vidulich, Michael A.; Nelson, W. Todd; Bolia, Robert S,; Guilliams, Nicole M.; McLaughlin, Annie B.; Donnelly, Brian P. "An Evaluation of Speech Controls for AWACS Weapons Directors." May 2004.

[71] Waller, M. C. and Lohr, G. W. "A Piloted Simulation Study of Data Link ATC Message Exchange." NASA Technical Paper 2859. Hampton, VA, 1989.

[72] Wickens, Christopher D.; Mavor, Anne S.; Parasuraman, Raja and McGee, James P. *The Future of Air Traffic Control: Human Operators and Automation.*, Panel on Human Factors in Air Traffic Control Automation, National Research Council, 1998.

[73] Wright, Helen. "Automatic type detection using suprasegmental features." *Proceedings on the International Conference on Spoken Language Processing*, 1998.

[74] Yangarber, Roman; Grishman, Ralph; Tapanainen, Pasi; Huttunen, Silja. "Automatic Acquisition of Domain Knowledge for Information Extraction", *Proceedings of the 18th International Conference on Computational Linguistics*, 2000.

[75] *http://www.tele.soumu.go.jp/e/system/satellit/air.htm.* May, 2004.

[76] *http://www.acarsonline.co.uk.* May, 2004.