**Table of Contents**

1. Identification and Significance of the Problem ........................................ 1

2. Benefits of a Voice Activated Cockpit ..................................................... 1

3. Perceived Automated Speech Recognition (ASR) Issues ......................... 4

4. Conclusions ............................................................................................ 6
Voice Activated Cockpits

1 Identification and Significance of the Problem

The primary objective for the crew of any flight is to get from point A to point B safely and economically. There have been many technological advances that have improved flight safety, including instrumentation facilitating flight without reference to the ground, radio communication, Ground Proximity Warning Systems (GPWS), weather radar, Traffic Collision Avoidance System (TCAS), Flight Management Systems (FMS), and so forth. Aircraft information technology generally provides the pilot more information with which to achieve better situational awareness. Multi-sensor navigational data coupled with improvements in display technology all provide better access to this information.

Unfortunately, each new technology intended to assist the pilot engenders additional complexity and increases pilot/computer interactions. The more disparate information sources there are in the cockpit, the more they must be visually scanned and their results mentally integrated into the pilots’ situational awareness. They typically require increased physical manipulation by the pilot of layered controls, menu-type access on displays, and multifunction buttons. In short, today’s pilot has become more of a “systems operator” and frequently spends more time manipulating his flight management system than he does actually manipulating the aircraft controls and looking out the window. This increased cockpit workload can ultimately distract from his real-time situational awareness. Such distraction has proven to be a significant factor in Controlled Flight into Terrain (CFIT) incidents.

2 Benefits of a Voice Activated Cockpit

Generally speaking, a pilot has three bidirectional channels for information flow – visual, manual, and auditory. He typically receives cockpit-generated information visually and responds/commands manually. His auditory channel is usually reserved for communications with ATC and/or his copilot and passengers. Under stressful flight conditions (e.g., abnormal or emergency flight situations, marginal VFR, shooting an instrument approach or looking for traffic while maneuvering on final), the pilot’s visual channel is maxed out, while his manual channel is moderately to heavily loaded, depending on the degree to which he is manually flying the aircraft or having to reprogram his FMS and/or
instrumentation for rapidly-changing flight conditions. During all of this, his auditory channel is usually only lightly loaded.

Data entry is a particular problem in a fast-moving vehicle like an aircraft. Keypads and keyboards common, are relatively easy to use, and are familiar to the generation that has grown up in the computer age. However, any type of keyboard is susceptible to input error and the keyboards and keypads typically found in an aircraft cockpit are smaller and more compressed than the full size keyboard found on an office desktop. Typing long strings of alphanumerics, especially during stressful and/or turbulent conditions, can easily lead to the entry of incorrect data which can go unidentified. Dials and switches are excellent quick, sequential data entry methods, but there are rather few data types that can be entered more efficiently using dials than with a keyboard. Just as keyboard entry requires close attention and hinders situational awareness, knobs and dials require even closer attention while sequencing through numbers or letters.

Even relatively straightforward tasks in general aviation require a number of appropriately-sequenced actions in order to execute them. For instance, to talk to a specific control tower, a GA pilot must divert his/her scan from traffic and instruments and at least one hand from the controls, find the appropriate frequency for that location either by spotting it buried somewhere on a paper chart or search for it in an FMS or GPS database, then dial the frequency into the radio, press the appropriate button to make it current, depress his PTT button, and – finally – speak.

One means of safely making cockpit interactions more efficient is obviously to exploit the pilot’s lightly-loaded auditory channel. Voice communication is very efficient. Suppose that a pilot could communicate with his cockpit like he does with his copilot. A simple “Google” internet search for “cockpit speech (or voice) recognition” will produce thousands of results, many of which are detailed studies about whether speech recognition might be a beneficial means for pilots to interact with aircraft systems. These studies go back more than twenty years, and with very few exceptions agree that if speech recognition were robust enough to deal with the challenging environment of an aircraft cockpit, then the technology would indeed be beneficial from both a safety and efficiency perspective.

A Voice-Activated Cockpit (VAC) could provide direct access to most system functions, even as the pilot maintains hands-on control of the aircraft. By “cutting out the middlemen” of button pushes and interpreting visual representations, the following safety and efficiency benefits now become possible:

- **Direct Aircraft Systems Queries** – Rather than step through menus to query specific aircraft systems or scan a specific instrument, a pilot could simply ask the aircraft what he wants to know, much as he would a copilot or flight engineer. For instance, “say remaining fuel” would cause a synthetic voice to report the fuel state.

- **Data Entry for FMS, Autopilot, Radio Frequencies** – Updating the flight profile in flight now becomes easier and safer, as there is far less likelihood of speaking the wrong lat/long or radio frequency than there is in inputting the incorrect data. (Naturally, we must assume that the VAC will “hear” what is said correctly – more on that later.)
• **Correlation of Unfamiliar Local Data** - ATC might issue a clearance to the spoken name Orlando, but the chart symbol (electronic and paper) may only show the seemingly-unrelated abbreviation “MCO”. With a VAC, the pilot need only repeat the name of the waypoint (or, even better, the VAC helpfully noticed it) and its underlying database will correlate the name with the chart symbol. When using electronic moving map displays, locating a specific waypoint or location could become as trivial as asking for it.

• **Glass Cockpit Configuration** – Today’s glass cockpits offer almost limitless configurations. A well-designed VAC would allow each pilot to configure the cockpit quickly to his or her preference by simply announcing himself when he took the left seat. Furthermore, different configurations could be defined for each flight modality and initiated as required via voice commands, e.g., the pilot may prefer a different cockpit configuration in cruise than he would on an IFR approach.

• **Electronic Flight Bag (EFB) Interaction** – Much like the direct aircraft system query, the pilot could simply ask the EFB to “Display the chart for NDB to ILS runway 27 Left at Orlando International”. This feature is particularly beneficial for example during a sudden change of landing runway when inside the terminal area. Requesting new approach data, navaid frequencies etc from the EFB serves only to distract the pilot during instrument cross checks in a critical phase of flight.

• **Checklist Assistant** – Here there are two application possibilities: 1) a single pilot aircraft operator might read through the checklist as the aircraft executes and confirms his instructions, or 2) the synthetic speech system leads the pilot through the checklist without the need to refer to its printed version. As the pilot reports compliance, the checklist assistant automatically moves to the next item. Again, these features would provide significant benefit in emergency situations and abnormal flight conditions.

• **Level and/or Heading Bust Monitoring** – A suitable VAC could monitor the pilot’s read-back of assigned ATC headings, altitudes, and altimeter instructions and compare the read-back against both what it heard ATC say as well as what the aircraft systems are reporting. For example, if the aircraft descent rate is not decreasing as the aircraft approaches the cleared level, the pilot might be alerted. A recent study by the UK CAA of level bust reports between 1990 and 1999 showed that by far the primary cause of level busts is “…a pilot failing to follow the level instructions even after a correct read-back.”

• **Memo Creation** – In a VAC, a pilot might record memos correlated to the position and route of the aircraft for later playback, e.g., “commence descent to flight level five zero on reaching “Ottringham”. The same feature can be used to simply record notes for the after flight report, e.g., “inform maintenance of slightly high temperature on engine 2.” The key here is that the VAC correlates pilot voiced instructions with aircraft state, before, during, and after the flight.
Perceived Automated Speech Recognition (ASR) Issues

Naturally, these VAC capabilities require sufficient accuracy of automated speech recognition (ASR), something many observers may dispute. However, Adacel believes that ASR technology has evolved to a level in which it is entirely practical as a cockpit interface and can overcome the technical challenges that the cockpit environment presents, namely:

- High noise levels
- A multitude of operator accents
- Changes in the speaker’s voice due to illnesses, air pressure, vibration, G forces, acceleration/deceleration
- Limited command sets because of the high level of hardware resources previously required to process speech
- The need to memorize and speak a limited set of commands
- The need to speak slowly, one word at a time
- The inability to differentiate between words or phrases intended as a command and those that are part of a conversation
- The need to train the system to recognize each operator’s voice patterns
- The difference between the printed word and the myriad ways in which the word or phrase spoken, e.g. “descend and maintain” may be spoken “descend-n-maintain,” or the word “route” can be correctly pronounced as either “root” or “rout.”

The perceived problems listed above have been resolved in the latest generation of ASR applications. The ASR system produced by Adacel for Lockheed Martin’s F-35 Joint Strike Fighter (JSF) is meeting seemingly impossible performance requirements:

- The system is speaker-independent and does not require any user to “train” the recognition system to recognize his or her voice characteristics
- The system must achieve a word recognition rate of 98% in high noise (up to 120Db) and up to 6G loading
- It must permit the chaining of up to three commands in a single utterance and allow correction of a misspoken command in the same utterance
- All of these performance requirements must be met within a 10% time-slice of a 750Mhz Power PC processor and within 10 Mb of memory

The following list includes examples of a few of the cockpit functions from around 100 commands (and millions of word permutations) that the JSF ASR system is being used to control:

- ILS and TACAN Control
- Communications systems control, frequency selection
- Transponder entry and control
- Targeting and Assignment
- Waypoint entry
- Steer point entry
- Autopilot control
- TFLIR Control
- Wing Fold & Launch Bar control
- Display management
- System information requests
- Command Macros
- Lights control

In addition to the system being deployed on the JSF, Adacel has developed and delivered to international customers (USA, Italy, Brazil) at over 130 locations worldwide, a speaker-independent ASR system that is being used as the sole human/machine interface in air traffic control (ATC) training simulations. This training system is remarkable in that it achieves very high recognition rates (98%) with a very large command set (literally billions of word combinations) and in challenging environmental conditions and operator profiles. The system that has been delivered to the United States Air Force, for example, has 5 operators and/or controllers, 15 computers, 6 projectors and a high capacity (i.e., noisy) air conditioning system within a room that measures only 6x5x3 m2.

Not only does the room present a challenge to the ASR system, so does the way in which the system is used. Air traffic controllers frequently have their frequencies linked to a loud speaker system in the tower cab. Multiple frequencies can be heard over the loudspeakers while the controllers are often also talking amongst themselves in raised voices as they negotiate and coordinate with each other to ensure safe movement of the traffic.

The ASR component of Adacel’s ATC simulator runs on a single Pentium IV PC. It accepts 4 simultaneous operator inputs, with each operator selecting among four discrete frequencies. Each speaker may link up to 8 complex multi-word commands in a single transmission (utterance) and also permits the correcting of those commands during the same transmission, e.g.,

“Eagle 1, Eglin Tower. Information Charlie, Altimeter 2 9 9 4, correction altimeter 2 9 9 2. Enter left downwind runway 2 7 left. Number two to follow the C one thirty at 5 mile final. Report left base, Traffic is F sixteen approaching left downwind.”

The command set (or grammar file) used in the ATC simulator is extremely large. Starting with a base set of commands of approximately 450, after considering the variations in each command, e.g., the taxi command has 9 variations (Eagle 1, runway 2 7 right, taxi) and the traffic information command has over 3000 variations (Eagle 1, traffic is an F fifteen approaching 3 mile final runway 2 7 left), the system has to deal with recogniz-
ing what was said from around 500,000 command variations. Given that we can link up to 8 commands together, the number of permutations of commands numbers in the many multiples of trillions.

4 Conclusions

Looking at the modern aircraft cockpit, it is easy to see how a pilot can be distracted by the task of “systems management”. The Airbus A380, for example, has a multitude of LCD displays, numerous gauges, desk space for full size computer keyboards and hundreds of buttons, dials, switches and knobs, some of which are rarely used.

![Airbus A380 Cockpit](image)

Airbus A380 Cockpit

Human factors specialists working on new aircraft cockpits such as the A380 are trying to produce interfaces that are intuitive and easy to use. But the sheer number of tasks that may be executed by the flight crew combined with the restricted amount of cockpit real estate available to display the information, as well as the criticality of aircraft weight as part of the aircraft design criteria, will always result in a cockpit environment that is suboptimal (human factors wise) and promotes more and more heads-down activity. An aircraft in which the flight crew can concentrate on flying the aircraft and ensure that the pilots gain and maintain situational awareness will always be a safer aircraft.

Adacel research scientists and human factors experts believe that speaking to the cockpit as a method of system management can become an effective interaction method (accepting that ASR works), since speaking is how we primarily communicate with each other. ASR in general, and aided and abetted by Adacel's proprietary heuristic technology, has advanced significantly in recent years and is now ideally suited to this application. Many hundreds of millions of dollars are being and will continue to be spent on ASR R&D, thereby ensuring that the performance of the VAC system will continue to improve at an exponential rate.
We believe that talking to aircraft systems will ultimately become second nature. Indeed, in his Keynote speech at SpeechTEK 2005, Bill Gates predicted that by 2011, the quality of ASR will catch up to humans. If he is right, in five short years, a “Star Trek” style of voice-activated cockpit should emerge.