

The Benefits of a Speech Recognition Enabled Cockpit

by

Adacel Systems Inc

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. 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), 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 adds additional complexity. 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.

Benefits of a Voice Activated Cockpit

Generally speaking, a pilot has three 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), 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. During all of this, his auditory channel is relatively lightly loaded.

Data entry is a particular problem in aircraft. Keypads and keyboards are susceptible to input error and the keypads typically found in an aircraft cockpit are more compressed than the keyboard found in an office. Typing long strings of alphanumeric, 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 just as keyboard entry requires close attention, 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 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. An internet search for "cockpit speech" 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 with very few exceptions agree that if speech recognition were robust enough to deal with the 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. The following are a few examples of 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. 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.
- **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 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.

- **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." <http://www.caa.co.uk/docs/33/CAP701.PDF>

Perceived Automated Speech Recognition (ASR) Issues

Naturally, these capabilities require a reliable and accurate automated speech recognition (ASR), something many observers dispute is possible. 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 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 pronounced as "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

In addition to the system being deployed on the JSF, Adacel has developed and delivered to international customers (USA, Italy, Brazil, Canada), a speaker-independent ASR system in air traffic control (ATC) training simulations. This system is remarkable in that it achieves very high recognition rates (98%) with literally billions of word combinations and in challenging environmental conditions and operator profiles.

The way in which the system is used presents a challenge to the ASR. Air traffic controllers often have their frequencies linked to a speaker 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 interact with each other to ensure safe movement of the traffic.

The ASR component of Adacel's ATC simulator accepts multiple simultaneous operator inputs, with each operator selecting among multiple frequencies. Each speaker may link up to 8 complex multi-word commands in a single transmission with the capability of correcting 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."

Conclusions

Looking at the modern aircraft cockpit, it is easy to see the complex workload of a modern pilot". 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

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 sub optimal (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, 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 by the industry, 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.