DOT/FAA/TC-24/25

Federal Aviation Administration
William J. Hughes Technical Center
Air Traffic Systems Test & Evaluation
Services Division
Atlantic City International Airport
New Jersey 08405

Computer-Automated Simulation Pilot (CASPER):

Integrating Automated Speech Recognition into Federal Aviation Administration (FAA) Simulation Systems

October, 2024

Final report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Computer-Automated Simulation Pilot (CASPER): Integrating Automated Speech Recognition into FAA Simulation Systems		5. Report Date October 2024 6. Performing Organization Code ANG-E5B
7. Author(s) Nelson Brown & Jonathan R. Rein		8. Performing Organization Report No.
9. Performing Organization Name and Address Federal Aviation Administration, Human Systems Integration Branch (ANG-E5B) William J. Hughes Technical Center for Advanced Aerospace Atlantic City International Airport, NJ 08405		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address		13. Type of Report and Period Covered 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The Federal Aviation Administration (FAA) has developed the Computer-Automated Simulation Pilot (CASPER) to enhance the realism and efficiency of air traffic control (ATC) simulations. Integrating automated speech recognition (ASR) technology, CASPER improves human-in-the-loop (HITL) simulations by reducing costs and addressing controller availability and resource optimization challenges. Its modular design and seamless integration with existing FAA systems, such as the Target Generation Facility (TGF) and Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE), allow it to support tower, terminal, and enroute operations.

CASPER operates in both full-response and suggestion modes, providing a versatile tool that enhances simulation fidelity and reduces the demand for human simulation pilots in lower-priority tasks. Demonstrations confirmed CASPER's accuracy and effectiveness, delivering high levels of immersion for controllers. CASPER also has potential applications in ATC training. Ongoing development in collaboration with the William J. Hughes Technical Center and Mike Monroney Aeronautical Center highlights CASPER's potential as a cost-effective and customizable platform for advancing ATC practices. Its adaptability complements commercial off-the-shelf (COTS) solutions, providing a strategic tool for enhancing simulations and training. By addressing current challenges and aligning with FAA goals, CASPER offers valuable capabilities for the future of air traffic management.

17. Key Words Air traffic control, modeling and simulation, automated speech recognition, artificial intelligence	18. Distribution Statement This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at https://rosap.ntl.bts.gov.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price

Contents

Figures

Acronyms

Acronym	Definition
AMQ	ActiveMQ
ASR	automated speech recognition
ATC	air traffic control
CASPER	Computer-automated Simulation Pilot
COTS	commercial off-the-shelf
DESIREE	Distributed Environment for Simulation, Rapid Engineering, and Experimentation
DIS	distributed interactive simulation
FAA	Federal Aviation Administration
HITL	human-in-the-loop
SME	subject matter expert
STARS	Standard Terminal Automation Replacement System
TCP	transmission control protocol
TGF	Target Generation Facility
TRACON	terminal radar approach control

Executive summary

The Federal Aviation Administration (FAA) has developed the Computer-Automated Simulation Pilot (CASPER) to enhance the efficiency, realism, and cost-effectiveness of air traffic control (ATC) simulations. CASPER integrates automated speech recognition (ASR) technology to improve human-in-the-loop (HITL) simulations, addressing challenges such as limited controller availability due to staffing constraints and the need for optimizing simulation resources. CASPER reduces the reliance on human simulation pilots, offers flexible deployment options, and supports a wide range of ATC environments, including tower, terminal, and enroute operations.

CASPER's modular architecture and compatibility with existing FAA systems, such as the Target Generation Facility (TGF) and the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE), enable it to operate in both full-response and suggestion modes. These modes provide a versatile toolset for varying simulation needs, enhancing both efficiency and realism. Successful demonstrations, including those conducted in simulated Denver and Northern California terminal radar approach control (TRACON) arrival scenarios, have validated CASPER's capabilities. ATC subject matter experts reported high levels of accuracy and immersion, demonstrating CASPER's ability to effectively automate pilot responses while maintaining operational standards.

Beyond its use in research and development, CASPER holds potential for ATC training environments. Its voice-based simulation interactions can be accessed remotely or integrated into web-based training modules, providing trainees with hands-on experience while supporting flexible, data-driven approaches to training. This adaptability is particularly valuable for accommodating trainees who cannot be physically present, offering the FAA new opportunities to refine training curricula based on real-time feedback and performance data.

CASPER offers a forward-looking, practical tool that enhances the FAA's ATC simulation and training capabilities. By improving efficiency, reducing operational costs, and supporting a more adaptable simulation environment, CASPER addresses current challenges while laying the groundwork for future advancements in air traffic management. Its evolution will continue to contribute to the FAA's mission, offering benefits to controllers, trainees, and researchers alike.

1 Introduction

1.1 Purpose and scope

The Federal Aviation Administration (FAA) conducts many simulations of air traffic control (ATC) operations for research, engineering, development, and training purposes at the William J. Hughes Technical Center for Advanced Aerospace and other facilities throughout the country. Many of these are human-in-the-loop (HITL) simulations, typically involving one or more controllers working traffic and communicating with simulation pilots. The Human Systems Integration Branch (ANG-E5B) has developed an extension to the Technical Center's existing ATC simulation system, designed to enhance the efficiency and cost-effectiveness of HITL simulations.

This new software—the Computer-Automated Simulation Pilot (CASPER)—integrates automated speech recognition (ASR) technology into the simulation pilot role, addressing the growing need to optimize resources in complex air traffic control scenarios.

The primary motivation behind this extension is to reduce the cost associated with conducting large and complicated simulation experiments. By leveraging a fully-automated simulation pilot, we can increase the number of shakedown events at a lower level of fidelity, reserving resources for fully staffed sectors during more critical data collection phases. This not only saves resources but also enhances the flexibility of our operations. The fully-automated simulation pilot could also be assigned to "ghost" or "feeder" sectors, which are controlled by air traffic controllers not directly involved in the study and therefore do not require the highest level of fidelity. For the high-fidelity cases where human pilots are required, the system can provide real-time, helpful suggestions to those pilots, improving their accuracy and efficiency.

This initiative aligns with the broader goals of the Human Systems Integration Branch, particularly in advancing future studies involving air traffic controllers. The added capability of using speech recognition in remote studies—conducted via web interfaces and teleconferencing software such as Zoom or Microsoft Teams—opens new possibilities for extending this technology into different environments, including potential training applications. Although not initially designed for direct training use, the system could provide valuable insights into how ASR can enhance training procedures, uncover drawbacks, and inspire innovative techniques to make training exercises more efficient.

1.2 Background

1.2.1 Target audience

This report is intended for stakeholders across the FAA, as well as external entities involved in ASR research within the industry or other government bodies, both domestic and international. The goal is to keep these stakeholders informed about our current efforts in integrating speech recognition into air traffic control simulations.

1.2.2 Current challenges

The development of this software is driven by the challenges associated with securing the participation of active air traffic controllers in our simulations. Due to travel constraints and staffing shortages, their availability is limited, making it crucial to maximize the effectiveness of the time they can dedicate to our simulations. To mitigate these challenges, we often rely on ATC subject matter experts (SMEs), typically retired controllers, during earlier stages of development and testing. By incorporating speech recognition and artificial intelligence, we aim to empower these ATC SMEs to engage in more realistic and immersive simulations, thereby enhancing the overall fidelity of our scenarios. Furthermore, pilot resources are often constrained due to limited budgets and a relatively small pool of well-trained pilot personnel. We designed CASPER to reduce the cognitive load on simulation pilots—who often manage multiple aircraft—by providing real-time suggestions through speech recognition, ultimately increasing the realism and responsiveness of the simulation.

1.2.3 Scope of the prototype

The prototype software developed under this initiative includes speech recognition models tailored for tower, terminal, and enroute operations. It is compatible with communication systems supporting the Distributed Interactive Simulation (DIS) protocols and has been successfully tested with the SimPhonics and ASTi Voisus communication platforms. Additionally, voice and audio data can be integrated via a web interface, and the system is capable of operating with teleconferencing software such as Zoom and Microsoft Teams. The prototype's broad applicability allows it to be used across various ATC scenarios that utilize the Target Generation Facility (TGF) traffic simulation system. Integration with the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) has enabled more rapid prototyping of new user interface enhancements or procedural changes in high-fidelity ATC simulations by providing closed-loop controller-pilot communications within the DESIREE interface. The prototype has been tested with the AgentFly system—an artificial intelligence agent-based controller used in ghost and feeder sectors that employs realistic human performance metrics to simulate air traffic controller behavior—to demonstrate a maximally-automated HITL simulation.

1.3 Expected impact

The CASPER prototype system is expected to reduce costs and increase the number of HITL simulation events by providing an enhanced, flexible toolset. Whether deployed remotely in lower-fidelity simulations or used to augment realism in high-fidelity environments, this prototype represents an advancement in the simulation capabilities of the existing simulation systems for air traffic control.

2 System description

2.1 System overview

The CASPER prototype operates by capturing and processing controller audio data using the DIS protocol. The audio signal is processed and transcribed to text. The text is parsed into the semantics of the controller command and associated pilot response. A complete list of commands that CASPER recognizes is in the Supported Phraseology appendix. The current prototype has two operating modes:

- Full-response mode: CASPER injects the simulation pilot command directly into TGF, triggering an update in the aircraft state and trajectory, and sends the appropriate pilot readback audio data
- Suggestion mode: CASPER sends a suggestion simulation pilot command to TGF, which routes it to the corresponding simulation pilot workstation for human review

The system is designed to handle multiple concurrent air traffic control positions/frequencies and simulations, which are configured via an initialization message sent by the TGF simulation system. After initialization, the aircraft in each sector are regularly updated, typically once per second. This data, which includes aircraft metadata like callsigns and current routing, is combined with telemetry data to build a model of possible controller utterances. The availability of call signs within each sector is crucial for accurate transcription.

2.2 System architecture

The CASPER system is modular, with components communicating over a message queue. The key components include:

- Voice Bridge: A Java-based module that translates audio data between DIS protocol and other standard formats.
- Audio Clients: These are desktop or web applications written in Java or Node.js that handle incoming audio and playback responses from remote users who do not have access to the laboratory communication systems.
- CASPER Server: A long-running service written in Scala with the following components:
  - Simulation Coordinator: Receives initialization, status, and telemetry updates from the TGF simulation system. Transmits the corresponding simulation pilot commands to TGF for injection or suggestion to human simulation pilots.
  - Audio Handler: Receives and queues audio responses for all simulation sessions.
  - Language Model Service: Combines language model updated from the simulation coordinator and acoustic model inputs to transcribe spoken utterances, parse these transcriptions into ATC concepts/commands, and generate the corresponding pilot response text.
- Acoustic Model: Written in Python, this component processes audio data and converts it into character probabilities.
- Text-to-Speech System: CASPER uses a commercial off-the-shelf (COTS) product called ReadSpeaker, which provides five unique voices. A custom Java wrapper ensures uniformity and customization in voiced responses.

Figure 1 shows the overall system architecture, including CASPER's internal components and the external systems that it interfaces with.

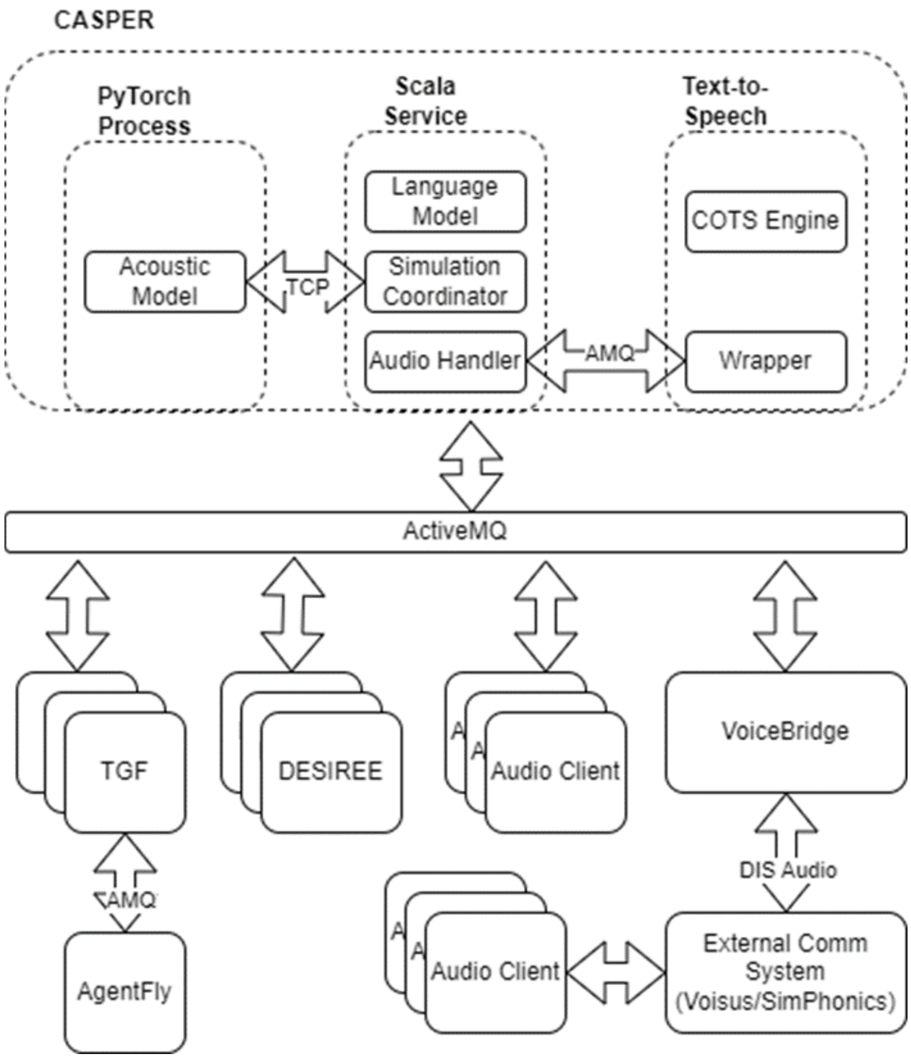

Figure 1. System architecture of CASPER

Each component operates independently within its own process, and they communicate via the message queue. This architecture allows for scalability, enabling the system to distribute processes across multiple servers if necessary to support large-scale simulations.

2.3 Integration with existing systems

CASPER is fully integrated with existing DIS communication systems, ensuring seamless interaction with the broader simulation environment. Figure 2 shows common data flows and the relationship between backend processes and the audio system, and the CASPER service components. TGF and external audio clients all interact with CASPER utilizing ActiveMQ. Internal messaging within the CASPER service uses Akka or TCP sockets between processes.

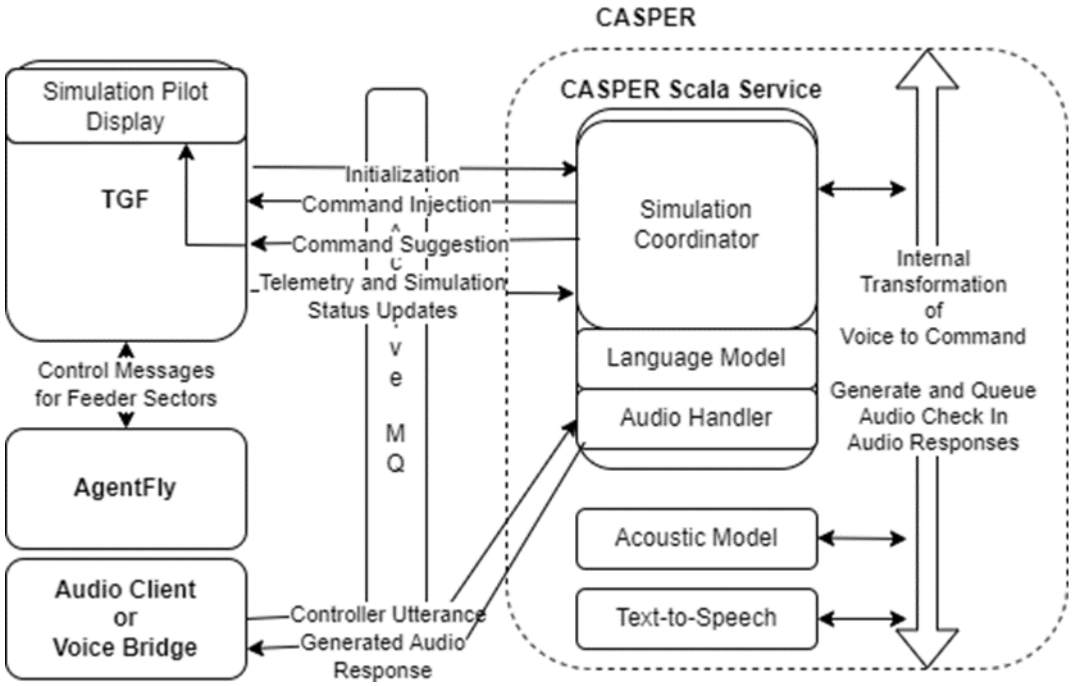

Figure 2. Data flows between CASPER and external systems

2.3.1 Target Generation Facility (TGF)

The TGF system reads current adaptation information for simulation sessions and generates an initialization message for CASPER, identifying each air traffic control participant, their audio identification, the desired CASPER mode for that frequency (full-response or suggestions only), and all relevant airspace adaptation data. Regular updates about aircraft and experiment states (e.g., start, pause, resume, termination) are also transmitted via the message queue, allowing CASPER to synchronize with the ongoing simulation. Figure 3 shows a simulation display of current aircraft driven by TGF to the left and on the right a CASPER user interface showing aircraft on frequency associated with the audio transmission, transcription, and simulation pilot command either injected into the simulation or suggested to a human simulation pilot operator.

Figure 3. Screenshot of TGF and CASPER.
Left: TGF's plan view display showing simulated targets and their data blocks. Right: CASPER's recognition table with a sampling of transcriptions.

2.3.2 DESIREE system

DESIREE provides high-fidelity replication of terminal, tower, and enroute automation environments. CASPER integrates with DESIREE via simulation data exchanged with TGF, which drives the simulated targets. A plugin developed for DESIREE allows CASPER to manage push-to-talk and audio information for remote HITL simulations conducted over web browsers and teleconferencing software. In its remote configuration, controllers can access DESIREE through a web browser and communicate with the CASPER pilot, enabling a fully dynamic simulation without requiring physical access to a controller workstation. Figure 4 shows a laboratory at the William J. Hughes Technical Center for Advanced Aerospace where the DESIREE system is driving several displays and manned positions of the air traffic control simulation session.

Figure 4. DESIREE system simulation and the William J. Hughes Technical Center for Advanced Aerospace

2.3.3 Communication systems

Audio data is collected through various systems that either directly support the DIS protocol or utilize switch panels to convert the analog signal used in operational environments. Custom bridging software converts the audio into a digital format and transmits it over a message queue using ActiveMQ, an open-source message broker application. The audio data is encoded using protocol buffers, which allows for consistent and efficient data transmission. The system supports multiple methods of audio input, including a web-based diagnostic page, teleconferencing audio connectors, and direct audio clients. The CASPER system has been tested with DIS traffic using SimPhonics and Voisus systems, both of which are in use at the William J. Hughes Technical Center for Advanced Aerospace. Figure 5 shows a typical audio client end-point used in the Voisus system.

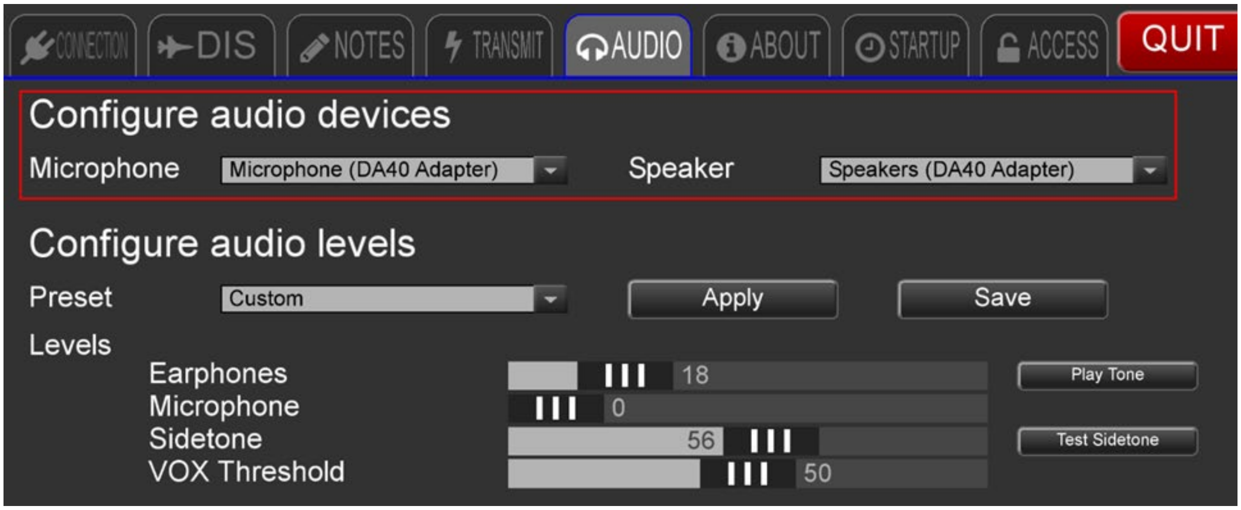

Figure 5. Voisus end-point client provided to simulation pilots

2.3.4 AgentFly system

The AgentFly system is a multi-agent airspace testbed that models the intentions and behavior of air traffic controllers or pilots. It is used to simulate feeder and adjoining sectors in enroute and terminal area sectors, reducing the need for ATC SMEs to staff adjacent sectors. Figure 6 shows a user interface that comes with AgentFly, which provides a diagnostic view of the model of all aircraft in the simulation, and which aircraft are controlled by a particular agent controller in a feeder sector.

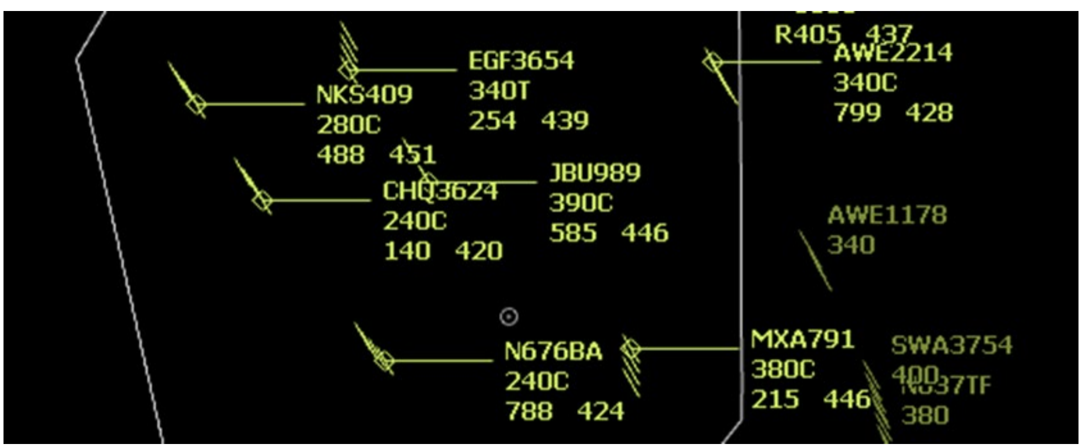

Figure 6. User interface representation of AgentFly controlled feeder sector

AgentFly also models cognitive load and workload measures based on traffic levels, adjusting controller behavior accordingly. Figure 7 shows some of the human performance metrics measured by AgentFly throughout a session. AgentFly does not directly interface with CASPER, but it has been used in tandem with CASPER during demonstrations to automate simulation activities more fully.

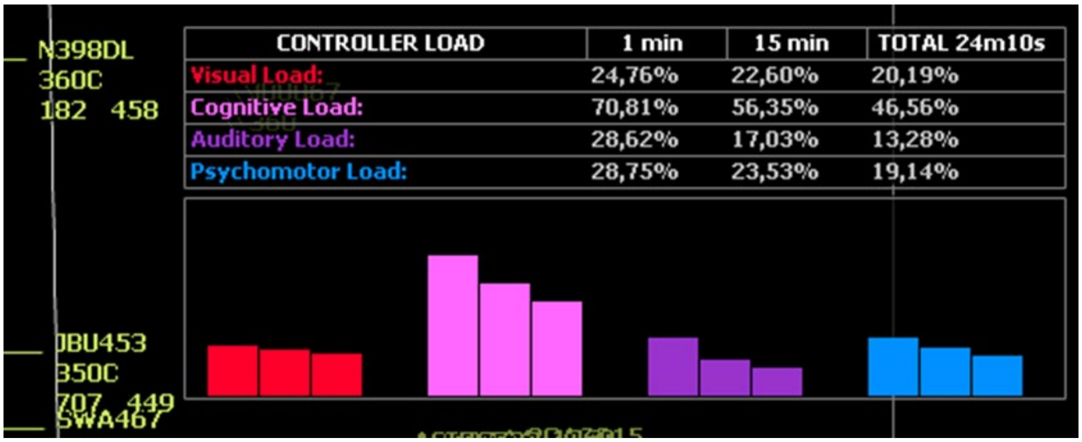

Figure 7. Human performance metrics calculated by AgentFly

2.4 Speech recognition

There are three critical components of the CASPER speech recognition system: the acoustic model, the language model, and the concept parser.

When audio is received, the signal data is resampled to 16 kilohertz. The signal data is then processed through an acoustic model. Provided signal data inputs, the acoustic model outputs a probability distribution over English characters at discrete time steps over the duration of the recording. The model was developed using the HuggingFace/PyTorch implementation of the Wav2Vec2 architecture (Baeyski, Zhou, Mohamed, & Auli, 2020). The model was initialized using publicly-available weights that were obtained by unsupervised training on the Librispeech dataset of recorded audiobooks (Panayotov, Chen, Povey, & Khudanpur, 2015). The model was fine-tuned using approximately 20 hours of transcribed recordings of air traffic controller speech, using the connectionist temporal classification method (Graves, Fernandez, Gomez, & Schmidhuber, 2006). The training corpus contained the commonly-used NIST Air Traffic Control Complete corpus (Godfrey, 1994) and more recent operational air traffic recordings that have not been released to the public.

The language model represents the probability of word sequences. CASPER's language model service builds a probabilistic directed graph of word-to-word transitions based on standard ATC controller phraseology and its variants. This graph is customized for the airspace of an individual sector and updated based on the set of aircraft on the controller's frequency. The graph is converted to a standard N-gram language model, which is essentially a hash table with short word sequences (typically two words) as keys and distributions over subsequent words as values. Given the character probabilities output from the acoustic model and the word sequence probabilities from the language model, we use a beam search decoder to identify the most likely transcriptions.

The concept parser takes a string of transcribed text and converts it into air traffic concepts. It uses parsing combinators (similar to nested regular expressions) built from common ATC phraseology and airspace adaptation (the relevant frequencies, fixes, runways, etc. for the sector). It extracts the aircraft call sign using a process that is robust to truncations and performs some validity checking of commands against the aircraft trajectory, ultimately yielding a single transcription with the required information to convert to a TGF pilot command. These commands are injected into the running simulation, and depending on the mode, either a simulation pilot provides a response based on the recognized command, or the command is directly injected, and a response is generated by a text-to-speech system.

2.5 User interaction

Configuration of the CASPER system is primarily managed through the existing TGF software, which generates the necessary initialization messages. CASPER can be deployed as a service hosted within the Research and Development Human Factors Laboratory at the William J. Hughes Technical Center for Advanced Aerospace, or it can be containerized and deployed as a long-running service, capable of connecting to multiple instances of TGF.

A web-based user interface is available for troubleshooting, interacting with the audio system, viewing transcriptions, and managing readback data. The system also records a database of audio utterances collected during simulation sessions, and test scripts are used during development to extend the training corpus or troubleshoot transcription and command translation issues.

2.6 Data flow and synchronization

CASPER operates as a long-running service. Initialization, simulation status, and updates of aircraft used to update the language model are shown in Figure 8. An initialization message from the TGF system is provided which includes the current airspace information on airports, runways, sectors, frequencies, and other associated adaptation data. The initialization message also includes information about sector configuration and speaker identification used to tie audio utterance meta-data from DIS or direct audio clients to the associated airspace.

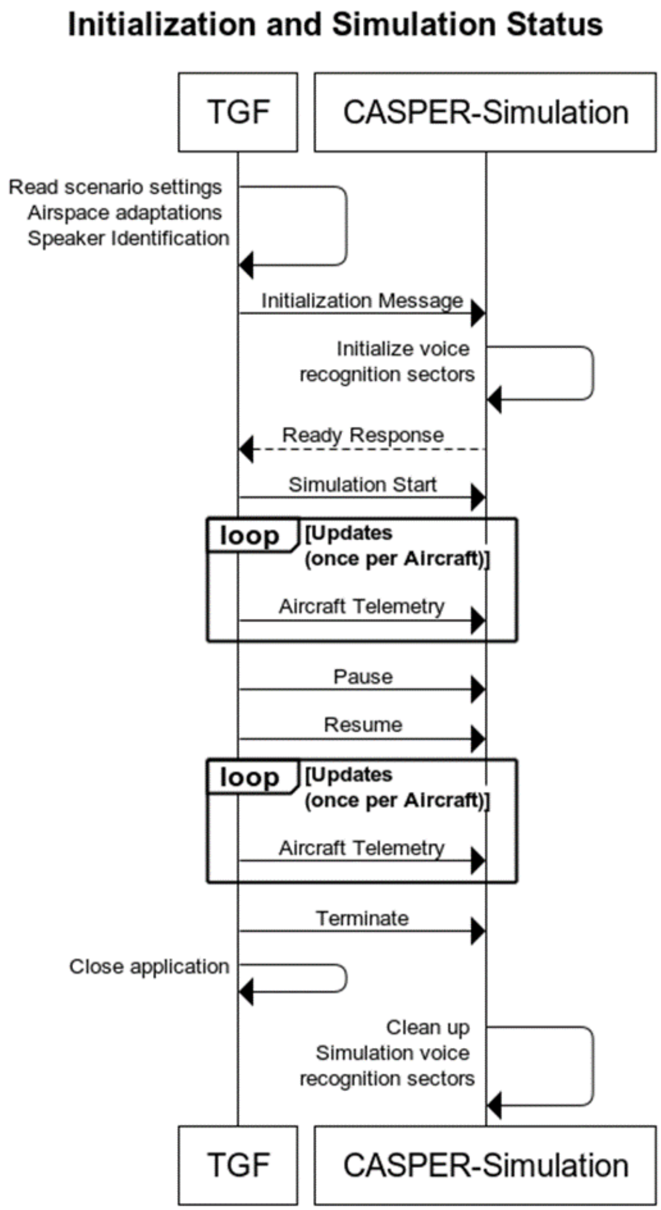

Figure 8. Sequence diagram of initialization and updates between CASPER and TGF

The data flow for a typical controller utterance when using CASPER is shown in Figure 9. After initialization, the DIS communication system will broadcast a DIS audio message, which will be passed to the VoiceBridge. The VoiceBridge will examine the DIS data and associate it to speaker identification, then convert that audio and transmit via ActiveMQ to the AudioHandler. This will associate the speaker identification to the appropriate sector and provide it to the acoustic model to generate character probabilities.

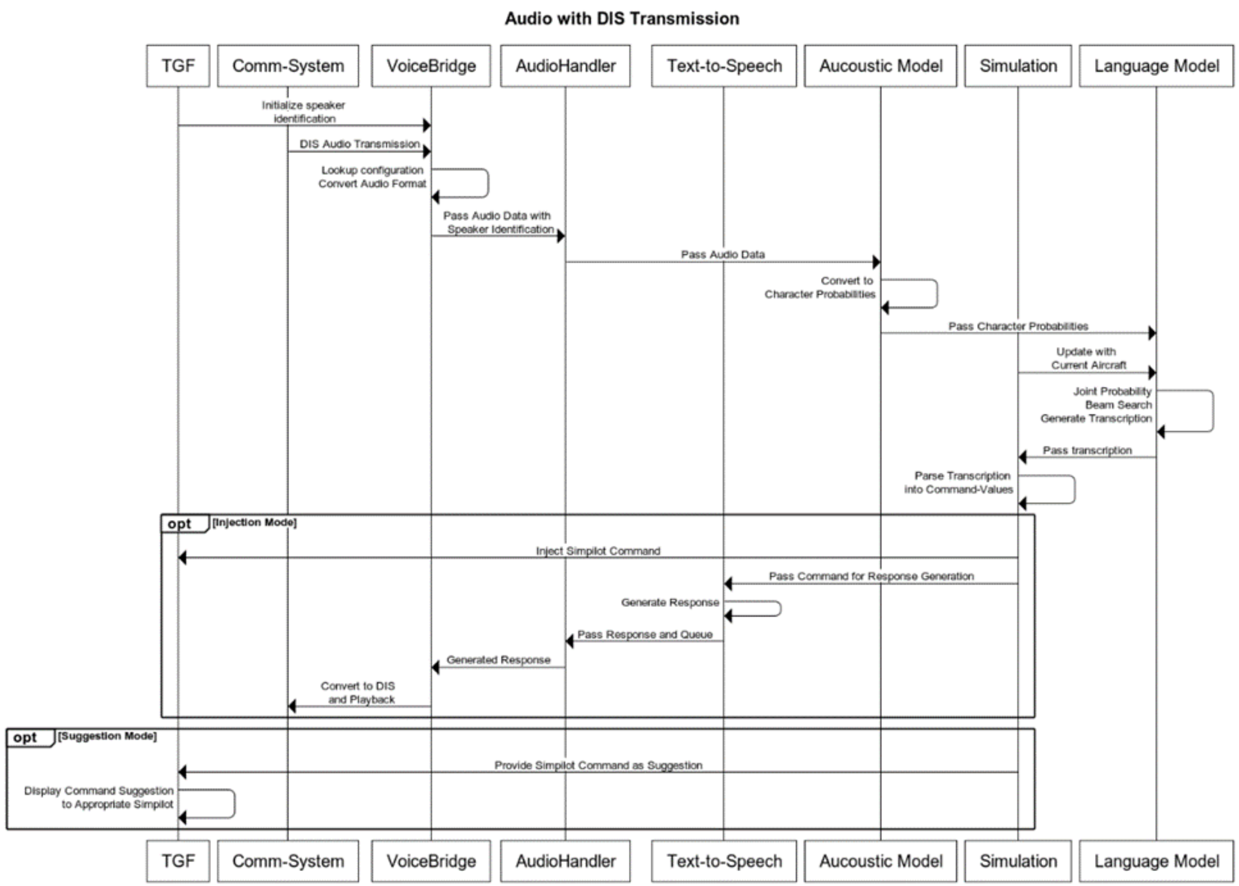

Figure 9. Sequence diagram of data flows from incoming DIS audio through CASPER response

The language model and simulation system will have received updates about the aircraft in the simulation. After receiving the character probabilities associated with the controlled sector, the system generates a transcription and performs a text-to-command translation as described in section 2.4 Speech Recognition. At this point, depending on the mode that the individual sector is configured for, either (1) the simulation pilot command will be directly injected into the scenario, and then an associated readback response will be generated and sent in the reverse path to the communication system endpoint, or (2) the simulation pilot command will be presented as a suggestion to a human simulation pilot operator who will accept or modify the command and issue an appropriate voiced response.

3 Performance and testing

3.1 Benchmarking on historical data

We assessed the accuracy of CASPER's automated speech recognition using a benchmark data set from a previous HITL simulation, which was run entirely with human simulation pilots. This previous study simulated a busy sector in the enroute environment, with typical commands across the range that CASPER was designed to handle (e.g. altitude, speed, heading, transfer of communications). We hand-coded the appropriate pilot response for all controller utterances for six 45-minute simulation runs. Out of the more than 500 transmissions, CASPER produced the correct pilot command 86% of the time. If we focus only on the 85% of transmissions where the controller used standard phraseology without errors or corrections, CASPER's accuracy was 96%. In practice, this latter level of performance can often be achieved with CASPER directly in the loop because ATC SMEs who know that they are talking to CASPER are more likely to stick to standard phraseology than when they are talking to real people, who they rightly expect to be more resilient than the current software. Either way, accuracy is substantially higher than the target level of 70% that we set at the outset of our development efforts.

3.2 Use in Conflict Alert human-in-the-loop simulation

In 2023, the Human Systems Integration Branch conducted a HITL simulation study, investigating the effects of proposed changes to the Standard Terminal Automation Replacement System (STARS) Conflict Alert system on controller workload and safety performance. Due to the research question being addressed, the traffic scenario simulated an extremely busy hour at Northern California TRACON, with a very large number of aircraft and a high level of complexity. Early phases of scenario development and shakedown required at least five simulation pilots to handle the quantity of aircraft, and pilots were frequently overwhelmed by the workload, sometimes operating hundreds of aircraft simultaneously. Although resource constraints did not allow for additional pilot personnel, CASPER had recently been developed and was ready for use. The research team conducted the later phases of scenario development and shakedown, as well as all eight weeks of official data collection, using CASPER in suggestion mode, providing recognized commands to pilots who approved/edited these commands and communicated with ATC SMEs over the radio.

Simulation pilots and ATC SMEs who participated in the study with and without CASPER reported a substantial improvement in their workload, accuracy, and response time when supported by CASPER suggestions. Without the automation support, pilots could not have effectively handled the number of aircraft. Using CASPER helped the project team achieve a high level of fidelity for the simulation without overextending resources.

3.3 Demonstration of full-response mode

In December 2023 and February 2024, Technical Center personnel demonstrated a maximally automated simulation using CASPER. TGF simulated hundreds of aircraft operating around the Denver airspace, with AgentFly simulating enroute controller behaviors, including clearing Denver arrivals on standard terminal arrival routes and handing them off to the Denver TRACON. Two ATC SMEs operated STARS positions in the Technical Center's Enterprise Capability Laboratory, working a feeder and final sector into Denver International Airport. Controllers communicated with CASPER through the operational voice communication equipment, issuing commands and confirming readbacks.

Collectively, the AgentFly enroute controllers and the real terminal controllers were able to complete a 30-minute Denver arrival scenario. The ATC SMEs reported that CASPER responded accurately and promptly to commands, and that they as controllers felt fully immersed in the simulation even though they were talking to a software-based pilot. By automating the enroute controller and simulation pilot roles, we were able to simulate a dynamic TRACON-based scenario with a fraction of the personnel typically required.

4 Applications, improvements, and future work

4.1 Possible applications for ATC training

In addition to research and development activities at the Technical Center, CASPER has the potential to be integrated into ATC training programs to simulate real-world air traffic control scenarios with higher fidelity. By incorporating CASPER and the TGF system into these training sessions, ATC training organizations can explore different methods and experiences for trainees against the baselines of existing products. This integration would be particularly valuable in the following application areas while fostering technical collaboration with the William J. Hughes Technical Center for Advanced Aerospace.

- **Controller Training**: CASPER's ability to simulate voice-based interactions with aircraft in various air traffic control environments—such as tower operations, terminal operations, and enroute operations—can provide trainees hands-on experience in managing complex airspace sectors.
- **Remote Simulation Exercises**: Trainees and instructors could participate in simulation exercises from various locations, with CASPER handling speech recognition and command processing in real time. This would be particularly useful in accommodating trainees who cannot physically attend sessions at the training facility.
- **Web-Based Training Modules**: CASPER could be incorporated into web-based training modules that allow trainees to practice their skills asynchronously. By interacting with the system via a web browser, trainees could receive real-time feedback on their performance, helping them improve their proficiency in air traffic control tasks.
- **Optimize Training Curricula**: Training organizations could analyze data collected from CASPER simulations to identify areas where trainees struggle or excel, allowing for more targeted adjustments to the training curricula.
- **Consultation on Training Tools**: Technical Center personnel may provide consultation services to other government agencies or industry partners interested in adopting similar speech recognition and simulation technologies in their training programs.

4.2 Enhancements and future development

CASPER, developed internally by the FAA, offers a potentially cost-effective implementation and broader deployment of automated simulation pilots for ATC simulation and training, avoiding the licensing and external contract costs typical of existing COTS products. This aspect is crucial for government operations where managing expenditures is a constant priority. Additionally, CASPER's flexibility allows it to be adapted for use across a range of air traffic control scenarios, including tower, terminal, and enroute operations. This adaptability makes CASPER a versatile tool for both training and operational testing. Considering its flexibility, CASPER could also augment COTS products currently in use

The suite of products developed and maintained at the William J Hughes Technical Center for Advanced Aerospace including TGF, DESIREE, and CASPER support rapid prototyping and iterative testing, essential for exploring the integration of new technologies and enhancing speech recognition in air traffic control practices. Its capacity for continuous updates and customization allows the FAA to implement changes swiftly based on real-time feedback and evolving training needs. In contrast, COTS systems may offer less flexibility for modifications outside their initial configurations. COTS systems are typically tethered to the continuous purchase and update cycles dictated by their broader customer base and the specific vendor's goals.

Utilizing CASPER could enhance the technical proficiency of FAA air traffic controllers through more dynamic training scenarios and promote an environment of innovation and continuous improvement. The strategic development of CASPER, particularly in collaboration with entities like the William J. Hughes Technical Center and Mike Monroney Aeronautical Center, aligns with long-term FAA initiatives preparing the groundwork for future advancements in air traffic management and safety.

5 Conclusions

The development and implementation of the Computer-Automated Simulation Pilot (CASPER provides the FAA capabilities to conduct more efficient and realistic air traffic control (ATC) simulations. By integrating automated speech recognition (ASR) technology, CASPER enhances the fidelity of human-in-the-loop (HITL) simulations while reducing operational costs and personnel requirements. This innovation addresses some of the challenges associated with limited controller and pilot availability and the need for optimizing simulation resources.

CASPER's modular design, flexible deployment options, and integration with existing FAA systems, such as the Target Generation Facility (TGF and Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE, enable it to support a wide range of ATC operations, including tower, terminal, and enroute environments. The system's ability to operate in full-response and suggestion modes provides a versatile toolset that can be adapted for different simulation needs, enhancing both the efficiency and realism of the simulations.

The successful demonstrations of CASPER in maximally automated scenarios, such as the Denver terminal radar approach control (TRACON) arrival simulations, have validated its capabilities, showing that automated simulation pilots can accurately respond to ATC commands and maintain high levels of immersion for human controllers. The feedback from ATC subject matter experts has been overwhelmingly positive, confirming CASPER's potential to reduce the need for human simulation pilots in lower-fidelity tasks while maintaining high operational standards.

Beyond research and development, CASPER holds promise for providing benefits to ATC training environments by providing realistic, voice-based simulation interactions that can be accessed remotely or integrated into web-based training modules. This adaptability supports a more flexible and inclusive approach to controller training, accommodating trainees who cannot be physically present and potentially allowing for more dynamic and data-driven training curricula.

The ongoing development and refinement of CASPER can provide a useful tool in collaboration with the William J. Hughes Technical Center for Advanced Aerospace and the Mike Monroney Aeronautical Center. CASPER's ability to provide rapid prototyping, iterative testing, and continuous updates may complement a strategy that relies on COTS solutions, offering the Federal Aviation Administration (FAA) a cost-effective and highly customizable platform to explore innovations in ATC practices, be they training or simulation.

CASPER offers a cost-effective approach to enhancing ATC simulation and training. By supporting current research and training needs, CASPER provides a practical tool for improving efficiency and reducing operational costs. As CASPER continues to evolve, it will enable more adaptable and realistic simulation environments, contributing to ongoing efforts in air traffic management.

6 References

Baeyski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). wav2ve 2.0: A framework for self-supervised learning of speech representations. *Advances in neural information processing systems*, pp. 12449-12460.

Godfrey, J. J. (1994). Air traffic control complete LDC94 . *Linguistic Data Consortium.* Philadelphia.

Graves, A., Fernandez, S., Gomez, F., & Schmidhuber, J. (2006). Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks. *Proceedings of the 23rd International Conference on Machine Learning*, (pp. 369-376).

Panayotov, V., Chen, G., Povey, D., & Khudanpur, S. (2015). Librespeech: An ASR corpus based on public domain audio books. *2015 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (pp. 5206-5210). IEEE.

7 Acknowledgements

CASPER could not have been developed and validated without the contributions of dozens of people who provided software development assistance, laboratory infrastructure support, and valuable subject matter insights. We tremendously appreciate all of the help from the following individuals: Vicki Ahlstrom, Kenneth Allendoerfer, Nigel Archibald, Stephanie Bell, Craig Bielek, Hareem Bokhari, John Bradley, Mike Capito, Shiraz Chaudhry, Dan Cordasco, Dan D'Ascenzo, Elisabeth Davis, John Dilks, Scott Doucett, Eric Einhaus, Jake Fowler, Brian Fox, Sheila Franklin-Smallwood, Shawn Frye, Mike Galusha, Adam Granich, Bill Griffith, Kevin Hallman, Laura Hamann, Patrick Hamill, Brian Hoy, Cliff Kaelin, Matt Kukorlo, Ed Little, Jonathan Lykens, Keith Maggio, Terence McClain, Tisa McKinley, Jim Merel, Rob Mullin, Eric Neiderman, Phat Ngo, Paula Nouragas, Mark Olsen, Nicole Peck, Mike Perseo, Bill Pfeiffer, Dana Picorale, Kia Price, Nicole Racine, Stan Rimdzius, Maryfaith Rodgers, Nick Roselli, Mike Ross, Rich Smail, Otto Smith, Lonnie Souder, Steve Souder, Rick Stoltzfus, Tim Swantek, Todd Truitt, Chow Truong, Mai Truong, Alex Tseng, Preston Wagner, Ben Willems, and Shelley Yak.

A Supported phraseology

The table below provides example phraseology of controller concepts/commands that CASPER recognizes, as well as the corresponding TGF simulation pilot command, where one exists. For commands with directionality (climb/descend, up/down, increase/reduce etc.), both directions are supported but the table generally includes only one direction for simplicity, unless the TGF commands differ. Some of the variants in phraseology are deviations from the standards prescribed in FAA Order 7110.65. CASPER “supports” these in the sense that it will recognize and respond accordingly; use of standard phraseology is recommended.

Controller Phraseology	Simulation pilot command
Altitude	
climb and maintain flight level two eight zero	A280
climb and maintain one four thousand	A140
climb at pilot's discretion and maintain one four thousand	A140
amend altitude maintain one four thousand	A140
climb and maintain one four thousand, expedite [your climb]	E140
expedite your climb through one four thousand	E140
climb and maintain VFR at or above four thousand five hundred	A45
Speed	
maintain two eight zero knots [or greater]	S280
speed two eight zero knots	S280
speed two eight zero	S280
reduce speed to two eight zero knots	S280
do not exceed two eight zero knots	S280
speed back to two eight zero	S280

Controller Phraseology	Simulation pilot command
reduce speed twenty knots	S-20
maintain present speed	S+0
maintain mach point seven six	M.76
reduce speed to mach point seven six	M.76
resume normal [air]speed	RS
normal speed	RS
cancel speed restrictions	RS
speed your discretion	RS
Heading	
fly heading two seven zero	H270
heading two seven zero	H270
turn left heading two seven zero	L270
turn right heading two seven zero	R270
turn left thirty degrees	G30
turn thirty degrees left	G30
turn thirty degrees right	D30
resume [own] navigation	RESUME
resume normal navigation	RESUME
proceed own navigation	RESUME
proceed on course	RESUME
[fly] present heading	H+0
hold present position	HPP

Controller Phraseology	Simulation pilot command
Communication Frequency	
new york center roger	**None**
denver approach radar contact	**None**
contact chicago center on one two seven point seven seven	CC 127.775
contact chicago one three three point three	CC 133.300
contact center one three three point three	CC 133.300
monitor tower one two three point four five	MF 123.450
ident	ID
squawk one two three four	SQ 1234
squawk VFR	SQ 1200
frequency change approved	SQ 1200
change to advisory frequency approved	SQ 1200
radar service terminated	SQ 1200
Routing	
cleared direct bogart (fix ID = BOGRT)	DIR BOGRT
cleared direct bravo charlie delta	DIR BCD
proceed direct bogart	RF BOGRT
cleared direction destination	DIR KACY
climb via the adam two departure (SID ID = ADAMM2)	SID ADAMM2
devin one arrival (STAR ID = DEVYN1)	CSTAR DEVYN1
descend via the devin one arrival	STAR DEVYN1

Controller Phraseology	Simulation pilot command
Crossing Restrictions	
cross bogart at flight level two eight zero	CRS BOGRT A180
cross bogart at or above flight level two eight zero	CRS BOGRT AOA180
cross bogart at or below flight level two eight zero	CRS BOGRT AOB180
cross bogart at two five zero knots	CRS BOGRT S250
cross bogart at flight level one eight zero at two five size knots	CRS BOGRT A180 S250
Approach	
[expect] runway two three right	RWY 23R
cleared for approach	CLA
cleared for ILS approach	CLA
cleared for ILS runway two three right approach	CLA 23R
cleared visual approach runway two three right	CLA 23R
cleared RNAV yankee two three right approach	RWY 23R; CLR Y
make straight in	CLA
straight in approved runway two three right	CLA 23R
start your VFR descent now	CLA
intercept the runway two three right localizer	RWY 23R; LOC
go around	MA
cleared to land	CLA
Ground	
runway two three right taxi via bravo charlie delta	TAXI 23R VIA B.C.D
hold position	STOP
resume taxiing	RT

Controller Phraseology	Simulation pilot command
hold short of runway two three right	HS 23R
cross runway two three right	CR
cross taxiway bravo	CT
lineup and wait	LW
cleared for takeoff	TO
Compound Commands	
fly heading three two zero and reduce speed two five zero knots	H320; S250
fly heading three two zero then reduce speed two five zero knots	H320 THEN S250

Environment	
traffic twelve o'clock five miles	**None**
traffic no factor	**None**
wind calm	**None**
wind three six zero at one two	**None**
altimeter two niner niner two	**None**