Human Systems Integration Approach in Implementing Voice-Control of Future Spacecraft Systems

George Salazar
NASA Johnson Space Center
2101 NASA Parkway
Houston, Texas 77058
281-483-0162
george.a.salazar@nasa.gov

Abstract—**Crewed spacecraft and habitats of the future will require more automation and autonomy to support complex missions. However, these complex systems require a more efficient command and control input method. Speech recognition along with visual or auditory feedback is an alternative, providing an extra pair of hands and eyes for the crew. Yet, speech recognition demands a highly integrated development approach to ensure a successful system implementation. To ensure the voice control application is developed correctly will require a Human Systems Integration (HSI) approach. This paper provides an insight into the development of a speech/voice control application of a spacecraft system that encompasses automation and autonomy through an HSI approach. Results of the voice control experiment of the Space Shuttle camera system are provided as lessons learned about voice control on a spacecraft. Limitations and challenges of the technology are addressed as well as how HSI can help develop these types of voice control command and control systems.**

TABLE OF CONTENTS

1. INTRODUCTION

It has been nearly 50 years since the last man walked on the Moon. The Artemis program sets the plan for returning to the Moon with the end goal of a sustained lunar exploration to learn how to live on another planet before going to Mars [1]. Mission to Mars and beyond round-trip communication delays between Earth and the spacecraft will be a challenge in terms of monitor and control of deep space assets such as vehicles/habitats. As an example of required monitoring and control, the International Space Station (ISS) missions involve over more than 100 Mission Control Center (MCC) flight controllers, mission evaluation room engineers, and Russian support flight controllers as well as engineering specialists to help astronauts operate the ISS. The ISS is a complex vehicle with more than 350 thousand sensors and more than 100 critical computers to monitor and operate. Future space systems going to Mars or beyond will be equally or more complex. Yet, the crew may be as small as four and cannot depend on MCC for immediate help during off-nominal and emergency conditions.

The communication delays will demand more onboard vehicle MCC capabilities through intelligent systems for maintenance and control. Hence, deep space missions to Mars and beyond will place challenging constraints on the crew that will demand an efficient and effective human-computer interaction (HCI) to control a highly complex vehicle/habitat system, including a spacesuit. The NASA Human Research Program (HRP) has identified inadequate HCI for future missions as one of the risks for deep space missions in part due to decrement in human performance (e.g., problem-solving and execution of procedures) [2] due to the deep space environment. Automation and autonomy (A&A) will be needed to help the small crew control a complex vehicle or extravehicular activity (EVA) spacewalk teaming with robots in orbit, on the Moon, or Mars. Voice control (VC) is a viable option as the technology permits the most common form of human communications—voice. If developed properly, a VC system could potentially help mitigate the HCI HRP risk.

The responsibility for goal-oriented human-machine interaction lies with the VC application dialogue. Machine-understanding dialogue is still a long way from capturing and understanding the speaker's intended meaning. Incorrectly recognized words must be treated cautiously as the system could branch into a wrong part of the application furthering the likelihood of more recognition errors. Implemented correctly, VC permits increased machine or system operator efficiency with a high level of recovery from recognition errors (REs). If incorrectly developed, the task workload can increase, resulting in potentially more REs, making the system unacceptable to the user [3].

Voice control of future space systems will demand a more comprehensive and methodical human-systems development approach that optimizes the effectiveness of the human-machine interaction focused on the user meeting the goals of the task. Human Systems Integration (HSI) is a process that optimizes the effectiveness of the human/machine interaction while considering safety and reducing life-cycle costs. It considers the human on par with the hardware and software in the given environment [4].

This paper provides guidance in developing a VC space system, primarily for command and control, using HSI as part of the systems engineering (SE) process. Whether it is an A&A or standalone application for space applications or other applications where the user interacts with the control of a system, the information provided in this paper applies. First, a high-level review of the evolution of human-computer interfaces over the past 60+ years is given. Then, a discussion of the constraints and challenges of implementing VC in a spacecraft going beyond low-Earth orbit is discussed. Next, a summary of the Space Shuttle Closed Circuit Television Camera system VC flight experiment is reviewed along with lessons learned from the first-ever control of a spacecraft subsystem by voice. Finally, a high-level approach in the development of a VC application utilizing HSI as part of the SE process is given.

2. Displays and Controls Evolution

Spaceflight missions have increased in complexity and so have the displays and controls. For example, as Fig. 1 shows, the complexity of displays and controls for spacecraft increased significantly from Mercury to the Orion program.

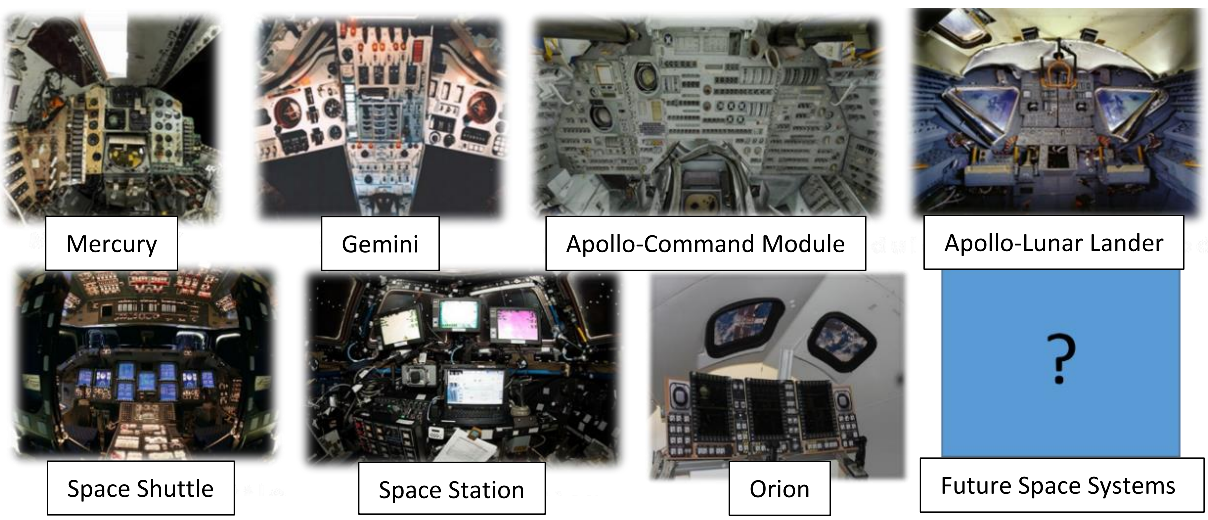


Figure 1. Evolution of Displays and Controls Systems (Photo Credit: NASA)

Mercury displays and controls had approximately 55 switches. Gemini program had more than 200 switches and controls. Apollo program became very complex, resulting in more than 700 displays and controls switches, and controls. The lunar lander had approximately 300 switches and controls. As we moved to Space Shuttle, the complexity of the vehicle resulted in more than 2500 displays, switches, and controls. International Space Station became more of a distributed control using laptops to control the space station over a data bus. Orion has evolved to use menu-driven displays to view more than 100 procedures and process pages and 50+ tactile switches for vehicle control. It is unclear what the implementation will be for future space systems as these systems will be more complex with the incorporation of automation and autonomy to control the vehicle, the habitat, or the spacesuit.

3. Voice Control Challenges

Though space operations have traditionally been performed through hardware components such as switches, keyboards, and now touchscreens, future missions will be much more complex, and yet the crews will be small. Therefore, VC would serve as an extra control input to the space system, particularly during simultaneous control and monitoring operations such as spacewalks and camera operations. From a potential benefits standpoint, VC would serve to provide [5]:

- Hands-free control
- Consistency of interfaces—human-computer interfaces use voice control
- The commonality of machine control usage throughout the Artemis space systems
- Machine control via voice rather than tactile switches or touchscreens without requiring diversion of visual attention from monitoring the task

Some of the key challenges are discussed in the following sections.

Technology Performance

Voice control technology has advanced since it was first introduced in the 1970s. Commercial industries such as banking, airline reservations, medical, and automotive (to name a few) are using speech control to help with customer inquiries or control of non-critical functions. Some industries such as Google use speech recognition as part of their search engines. Yet, despite the use of large speech recognition server farms, such as what Siri and Alexa use, performance has been mixed. Natural language processing has shown promise but still has issues recognizing a conversation. Factors such as prosody of the speech which indicates grammatical structures and stress of a word that suggests importance in the sentence have yet to be resolved. The ambiguity of the English language is also another challenge. The technology can recognize the words spoken but lacks an understanding of the meaning of the words. Homophones (words that sound the same but have a different meaning) can confuse the natural language processing application. Similarly, word boundary ambiguity recognition can pose speech recognition difficulties (e.g., recognize speech vs. wreak a nice beach) [6].

Despite the large computational assets, the technology still has challenges in terms of user acceptance. In one assessment related to digital assistants, Alexa beat out Siri and Google assistants. Yet, the decisive factor is the navigation of the queries by the user [7]. In addition, challenges with spoken word accuracy affected by background noise, accents, and human speech variability influenced by task loading and the user's psychological and physiological health still need to be resolved.

Intrinsic Speech Variations

A major challenge with VC is dealing with the intrinsic speech variations of speech production. As Fig. 2 shows, many variables affect speech production.

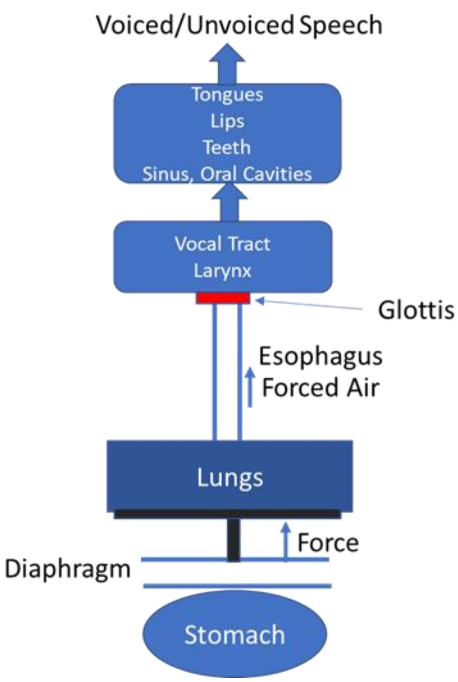


Figure 2. Simplified Voice Production Diagram
(Diagram Credit: Author)

The lungs, diaphragm, and stomach serve as the airflow voice production source that appears at the larynx. The larynx contains the vocal cords that control the airflow from the lungs to produce spoken words. This modulation creates a pseudo-periodic pressure wave impulse with a specific frequency determined by muscle tension force and the mass and length of the vocal tract. These factors vary by age and gender. In addition, the vocal tract cavities filter the airflow between the glottis and the lips. The lips, tongue, and teeth further modify the airflow to produce voiced (e.g., vowels) and unvoiced sounds-low energy (e.g., "S "as in stop) [8].

In addition to intrinsic speaker variations, the environment, physical, and cognitive stress in a space mission can further affect the speech process and recognition. For example, the microgravity environment causes no sinus drain, and micro-gravity causes muscular atrophy that could alter the speech process.

Acoustic Environment

Environmental noise plays a significant role in speech recognition performance in the presence of stationary and/or non-stationary noise. Technology does well when the signal-to-noise ratio (SNR) is better than 25 dB, such as in an office environment. However, as the SNR drops, so do recognition results [9]. The combination of enormous reverb such as in a spacesuit and both stationary and non-stationary noise (e.g., pumps kicking on and off) poses challenges for VC in space applications. The use of deep neural networks to predict the acoustic/noise environment has shown promising results [10]. However, note that these deep learning neural networks use General Purpose Graphics Processing Units (GPGPUs) that are not yet suitable for the deep space environment.

Spacecraft volume and materials, as well as layout, can affect the reverb of a speech signal, causing smearing of the spoken words [11]. Spacecraft systems typically have loud (> 85 dB Sound Pressure level) uplink voice messages from ground to space vehicle as well as caution and warning audio messages with a similar acoustic intensity that can leak into the microphone and corrupt speech recognition. A loud cabin environment can create a Lombard effect for the speaker causing the crewmember to speak loudly over the background chatter/noise [11] resulting in saturation and distortion of the input signal. EVA physical work strain on the crewmember can alter their breathing/speech. With background EVA noise, it can add to the speech recognition issue. The same can occur with tasks that have a high cognitive workload [12].

One study looked at Apollo 11 and understood how speech production changes in the space environment [13]. Acoustical features such as fundamental frequency and phoneme formant structure related to the speech production system were studied. It was noted that the combination of the changing environment and stress affects speech production. The conclusion was that speech technology must be adaptive to changing space mission phases in terms of acoustics and stress.

Speech Recognition Processing Performance and Galactic Radiation

Though speech recognition performance for Siris and Alexa is quite good, the commercial-off-the-shelf (COTS) technology is not easily usable for deep space systems where radiation galactic effects on electronics are important. Besides being huge systems and power-hungry, the system dialogue is designed for a search-type application rather than command and control. Also, the servers used for Alexa and Siris as well as IBM Watson are not designed to operate in the deep space radiation environment. As Fig. 3 shows, the Earth is protected by its magnetic field to deflect deep space high-energy photons and heavy ions. As we go beyond the magnetic boundaries, the radiation environment can have a significant impact on the electronics performance in terms of memory single-event upsets and single-event functional interrupts. This is significant as the constraints of the galactic radiation environment affect the processing capabilities of the hardware [14]. Components designed for deep-space radiation are roughly ten times slower and expensive.

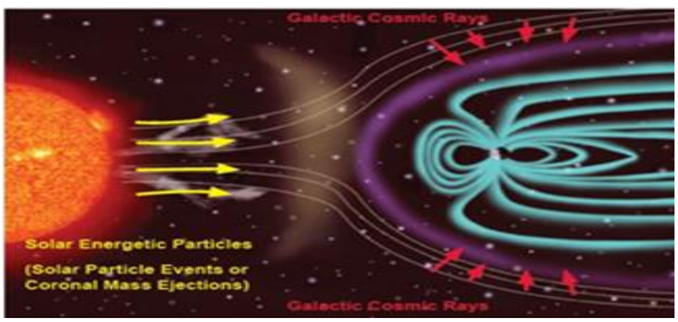


Figure 3. Earth's Magnetic Field (Credit: NASA)

Therefore, high-performance computing devices such as GPGPUs or Digital System Processing units will be difficult to find for high-performance computing (HPC) that meets the deep space radiation environment. For a deep space mission, the size of computing power will need to be assessed during the development phase. It is highly unlikely that the computing power Alexa, Watson, or Siri use will be able to port to a space environment application as these systems were designed for a large user population. Deep space missions will have a small crew size. Yet, some form of HPC will be needed to support future space missions—not just voice control but a limited artificial intelligence application.

4. SPACE SHUTTLE FLIGHT EXPERIMENTS

Two spaceflight voice control flight experiments were conducted on missions STS-41 (a 7-day mission) in October 1990 and STS-78 (an 18-day mission) in June/July 1996. NASA/Johnson Space Center Engineering Directorate developed a Voice Command System (VCS) flight experiment to evaluate speech recognition technology for space applications. The experiment would assess the operational effectiveness of using voice control to command a spacecraft system by voice. STS-41 was the first-ever demonstration in space of controlling a spacecraft system by voice [15].

Fig. 4 shows the photo of the VCS components and Fig. 5 the VCS architecture interfaced with the Closed-Circuit Television (CCTV) system. The Space Shuttle CCTV system comprised of a CCTV system controller that controlled monitor and camera selection, pan/tilt units,a video recorder, and the CCTV switch panel that sends commands to the CCTV controller.. Note from Fig. 5 that the VCS command interface control of the CCTV system was optically-isolated to the Space Shuttle CCTV switch panel command controls. The design permits sending discrete commands to the CCTV system either from the VCS or the Shuttle CCTV switch panel. The VCS optical-isolation was included as a safety feature in case the VCS failed to send a CCTV command (e.g., pan left). If that happened, the astronaut could use the CCTV switch panel to send the command(pan left switch on the CCTV switch panel) to the CCTV system. Or, if the VCS had major issues and had to be shut down, the astronauts could still use the CCTV switch panel to command the CCTV system.

Selection of the recognizer took approximately two years to converge on the best one. The speech recognizer chosen was a military speaker-dependent system that used speaker templates captured during training and stored in non-volatile memory. Each of the two astronauts assigned to the experiment had templates that were loaded into the recognizer memory based on stating who they were at power-up of the VCS.

Figure 4. Voice Command System Hardware Components (Credit: NASA)

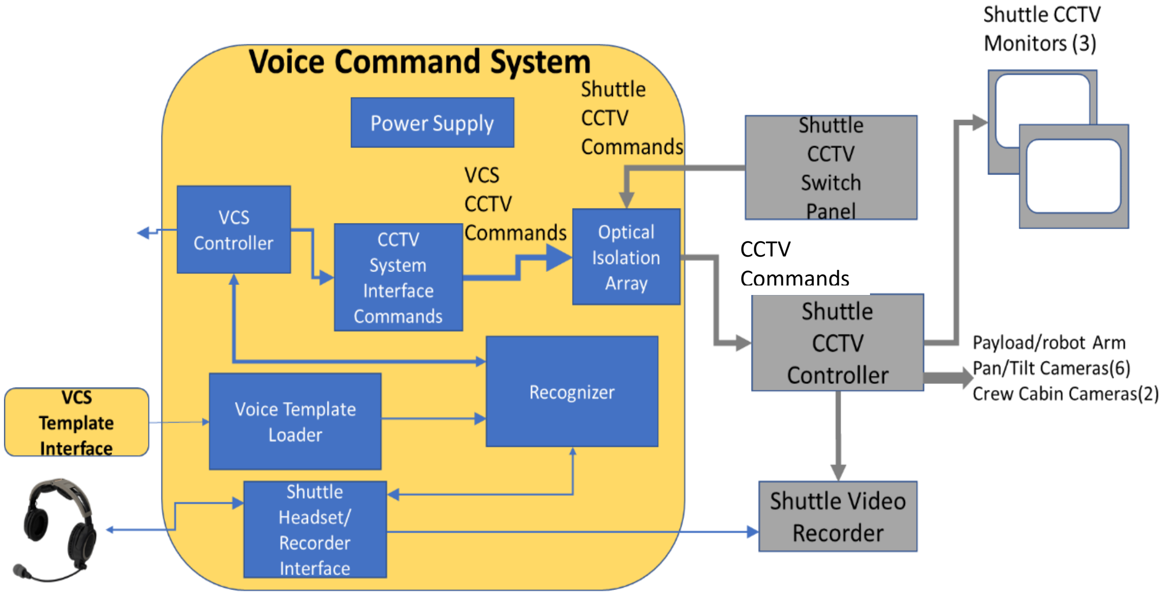


Figure 5. Voice Command System Architecture Interface to Shuttle CCTV System (Credit: Author)

The final speech templates were loaded into non-volatile memory in the last training session six weeks before the launch. The vocabulary was structured in a nodal/state machine fashion [15], requiring a transition word to be spoken to enable a small set of active command words for that node. For example, "activate" or "configure" would transition the astronaut to a node with expected command words the system recognized. An added feature was macro commanding, where one recognized command would initiate an automated execution of numerous commands to the CCTV system. For example, "Stow Cameras" would initiate a series of commands to stow all four payload bay cameras before returning to Earth.

Recognition accuracy scoring analysis for both astronauts showed that the number of words correctly identified dropped approximately 10% from the ground testing. However, the retrain feature that permitted on-orbit training of the command words showed an accuracy of both astronauts on average of 97%. Some difficulties were training the words as occasionally uplink voice calls from mission control would corrupt the training of the word requiring repeating, which was frustrating at times. Although the flight experiment did not capture voice samples to assess any changes in speech production, the flight showed the possible utility of using speech control in space.

The STS-78 mission used the same enclosure, but the electronics inside were redesigned to accommodate a COTS speech recognizer with adaptive recognition capabilities that continued to learn the user's voice the more it was used. The experiment vocabulary did not change significantly but added command words to accommodate the adaptive recognizer capabilities such as confidence check to tune the vocabulary [16]. In addition, to help the system adapt to the user, a query feature (Which of these top 3 words recognized was spoken.) was added during questionable recognition to adapt the word to how the word was spoken during real-time recognition. Also, a hardware feature was added to adjust the gain of the microphone input signal automatically based on recognition results.

This time, voice samples were captured to analyze the voice spectrum compared to ground testing. Analysis showed that there was a frequency shift in the vocal formant frequencies [16]. Like STS-41, STS-78 provided macro commanding. One observation in the use of voice control of a system is that it does not work when trying to use the technology like a mouse (point and click). One astronaut tried using the system for tracking a location on Earth with a payload pan/tilt camera with difficulty resulting in reverting to using the CCTV switch panel.

Overall, the adaption feature worked reasonably well, providing average recognition accuracy for both astronauts of 90% throughout the mission. However, the VCS had limited capabilities in adjusting for the placement of the headset microphone.

Several key Space Shuttle lessons learned (SSLL) from these experiments were:

1) Astronaut training time is a premium—they have many systems training for a mission. The project was given only eight 1-hour sessions for training the astronauts on the system. The system must be easy to use and operate. Therefore, the schedule should include a large amount of time in the early development phases to develop the design in terms of usability and training.
2) Feedback is important to know where they are in the command structure. The VCS had a one-line display, but it was not enough information to help astronauts understand where they were in the command structure. Lack of display feedback was noted during macro commanding where several commands were sent to the CCTV system to move the payload cameras to a certain view. They had to rely on visual viewing of the cameras pointing when done.
3) The development team should have at least one person that understands linguistics for vocabulary development. Word energy, number of syllables, and meaningful command words are important.
4) The selection of a recognizer requires a considerable effort in determining if it is suitable for the application. Vendors will advertise high recognition accuracy but fail to explain the conditions the scoring was based on.
5) Query checking regarding what word was said during actual commanding is an annoyance. Perhaps non-commanding applications will be best.
6) A fundamental principle of user acceptance is having a system that makes it easy to use and understand—even after months of not using the system. Sadly, the training schedule was too tight to iterate on the design to improve its usability. Future voice control system development efforts must include this in the timeline.
7) Speaker-dependent systems work best in the actual environment, but the number of command words needs to be kept to a minimum as it can be frustrating training command words in an actual mission.
8) Macro-commanding is useful if the machine has accuracy and repeatability that permits creating a plurality of commands that work every time (similar to automation). The Space Shuttle camera pan/tilt unit mechanism was not accurate in terms of the number of degrees panning and/or tilting per second. The final pan/tilt position would vary.
9) The system must be adaptive to the acoustic environment such as reverb or background noise. Uplink audio and caution/warning are very loud and can affect speech recognition. Spacesuit background noise is even more challenging to mitigate.
10) Dialogue design is critical to the successful implementation of voice control. Feedback of the recognizer on what word or phrase it thought it heard along with the status of the externally controlled system is important for situational awareness.
11) Engage human factors personnel early before the hardware is developed.
12) Mouse operations (point and click) are not suitable for voice control.

SSLL # 7 and 9 are more related to the technology than the other 10 SSLL that consider processes and procedures.

5. NASA DEVELOPMENT PROCESSES

To this point, the paper has focused on the challenges of voice control for a deep space mission as well as a discussion on Space Shuttle voice control flight experiments and some of the key lessons learned. Because of the highly human-machine interaction interface, the preferred approach in developing a voice control application is the use of HSI as part of the SE process. First, a high-level summary of the NASA SE and HSI process is given. Then, key methods during the SE and HSI process regarding voice control are discussed.

NASA Systems Engineering Engine (SEE)

Like many other companies and agencies, NASA uses a system engineering process to develop complex systems. Fig. 6 shows the NASA SEE used and the life-cycle reviews in the development of complex projects/programs. It is similar to the V-model (because it looks like a letter V) SE process described in the SE literature.

The left side of the SEE defines the processes for the definition and decomposition of the system. The right side represents the integration/verification/validation of the system. The center of the engine defines the technical management processes necessary to manage the development effort. Technical planning, requirements management, technical risk management, and interface management are key management areas of a program/project. Key life-cycle milestones shown are system requirements review (SRR) for all stakeholders to agree on the requirements, preliminary design review (PDR) to arrive at the best approach to solve the VC challenge, and critical design review (CDR) [17] to develop the actual flight system.

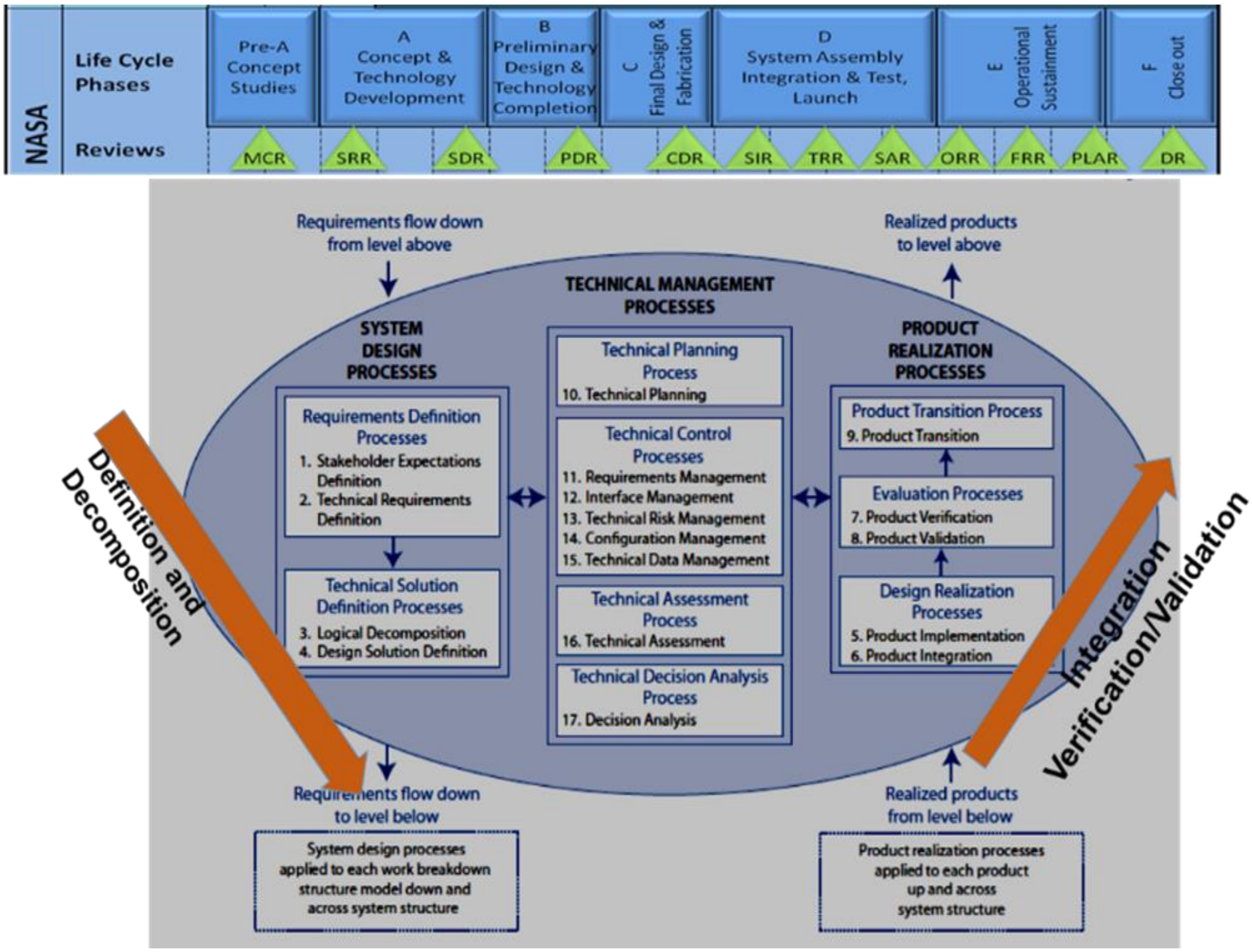


Figure 6. NASA Systems Engineering Engine (Credit: NASA)

Human Systems Integration

Since the 1960s, NASA has always employed human factors in system design to ensure the safety and protection of its spaceflight crews, focusing on human health and performance during the spacecraft/mission design and flight. With the increase in electronic component capabilities, complex software development, such as artificial intelligence and machine learning, and missions with increasing complexity, NASA is mandating applying HSI as part of the SE effort [18]. HSI goes beyond human factors engineering as it includes all domains that affect human performance. Each agency and company has its HSI domains that are critical to integrating the human with the hardware and software. Fig. 7 shows the HSI domains for NASA and a short definition.

Figure 7. NASA HSI Domains and Short Definitions (Credit: NASA)

A key takeaway from Fig. 7 is that the NASA HSI domains reflect all areas that affect the human element in a mission. Also, not shown is the double interaction that goes on as part of HSI—within the domains and between the different domains—to optimize the system/mission design. Like the SE process, HSI has a key process that integrates into the development process flow, the Human-Centered Design (HCD) process, that operates in conjunction with the SE process for the development of the hardware/software [19]. Fig. 8 shows the HCD as used by NASA. The process is iterative and is performed early and during final integration and verification activities. The next section addresses the details of the voice-control development process using HSI.

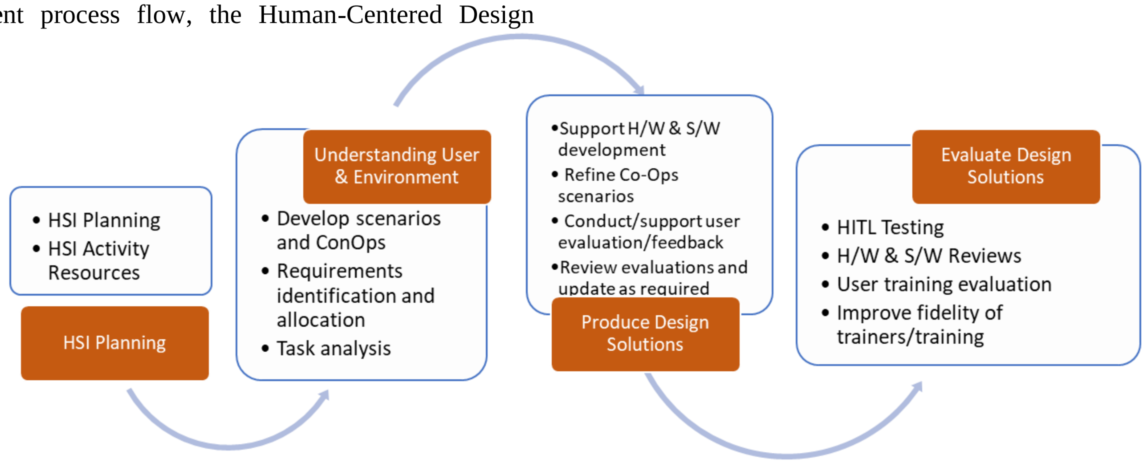


Figure 8. HSI Human-Centered Design Approach
(Adapted from NASA Human Integration Design Processes [19])

6. SYSTEM DEVELOPMENT APPROACH USING HSI

Referring back to Fig. 6, the definition and decomposition of SE activities are the critical part of the development effort. This is where requirements are developed, scenarios/tasks identified, speech recognition technology evaluated through prototype/evaluation test, and early human-in-the-loop (HITL) testing with the selected recognition system is performed. The preferred approach is to do HITL in the simulated environment to capture design shortcomings and mature the VC system application. If the VCS is part of a space mission design, as required by NASA Procedural Requirements (NPR) 7123.1c, the maturity of the technology should be at Technology Readiness Level (TRL) 6 [20], component/technology tested in the relevant environment.

For a VC application, reaching TRL 6 requires a considerable amount of planning, designing, and testing, especially if the application is mission-critical. Since a VC system could be risky in terms of technical performance, NASA could elect to first perform a feasibility study before committing the technology to a mission to ensure the development has a high degree of success. Regardless, the process described here can still be used. HSI domain experts related to the six NASA HSI domains, as shown in Fig. 7, should be involved at the onset of the development of the system. This is particularly important if voice control will be used on multi-vehicles and spacesuits to ensure commonality is applied of voice control application across all vehicles/habitats.

During definition and decomposition activities, an HSI practitioner or subject matter expert (SME) gets involved early to help plan the voice control development. The key HSI activity as part of SE is the HCD process. As shown in Fig. 8, the process helps ensure [21]:

- The task and environment are identified.
- The astronaut office/crew is involved throughout the design and development process.
- Design and requirement refinement occur through each design/analysis cycle of the prototype testing.

The development effort begins with the goal that has been defined, such as commonality of speech control with all command and control systems, or a specific task, such as control of robotic camera system.

Fig. 9 shows a notional high-level integrated SE and HSI process as it would apply to develop a VC system. HSI planning, along with SE technical planning, establishes the HSI plan and activities in developing the VC system. During stakeholders' expectations and technical requirements definition, HSI personnel become involved with HSI planning and HCD activities as part of the SE effort, requirements development, particularly about the human element, and allocation of requirements to the human and the hardware/software.

For A&A, requirements must include off-nominal situations when the user needs to take over control via voice control. This means that the A&A must provide information in terms of data information on the A&A status to the VC software and crew for situational awareness to ensure a clean takeover using voice. Also, requirements should include wireless applications where commands are sent over radio frequencies, Wi-fi, or Bluetooth. That interface must provide connectivity status to ensure the user knows that commands are being sent and the status received from the system under control.

Referring to Fig. 9, maturing the VC solution requires ensuring the task matches the human performance, the design matches the task, and the design matches the human to perform the task. HITL evaluations begin as early as possible in the life-cycle to flesh out technical issues with the human and the system. Usability and workload assessments under simulated worse case conditions that affect humans, such as the task's stress and background noise/reverb effects, help mature the design.

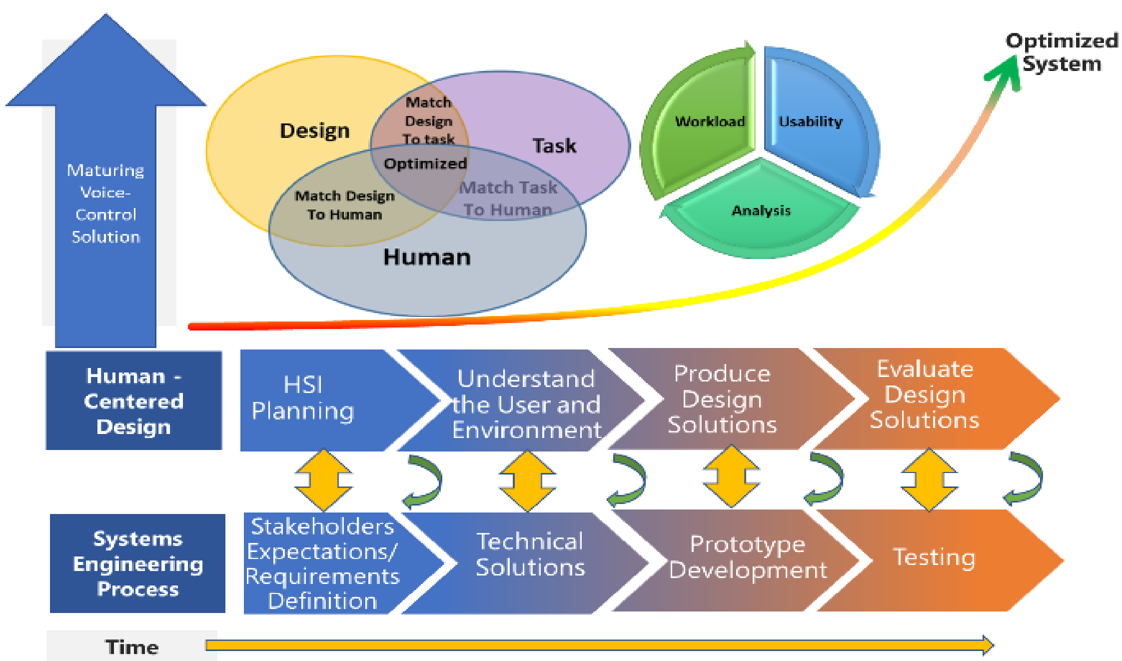


Figure 9. Notional Integrated Systems Engineering and HSI Process for Voice Control Development (Credit: Author)

Then, the Concept of Operations (ConOps) and task scenarios are defined using use cases, time sequence diagrams, or storyboarding. The ConOps should define not just the nominal but off-nominal as well to ensure safety is assessed for cases were operating a system could result in a hazard. Tasks are defined and analysis is performed to assess hazards or user errors. Finally, developmental HITL testing is performed, and the results are evaluated in terms of workload and usability of feedback to the requirements and design process. Note that the SE technical management process and the HSI management process collaborate In the areas such as requirements/risk management and decision analysis. This process is iterative until the measure of effectiveness and performance goals defined during requirements elicitation are achieved.

Fig. 10 shows some key factors to consider in the development of the VC system. Recognizer selection and hardware is primarily an SE responsibility working with the hardware/software development team. This would involve evaluating candidate recognizers in simulated environments using a preliminary vocabulary to assess how well it performs in various background noisy environments as well as understanding different accents from both men and women. Tests such as a confusability matrix shown in Table 1 could be used to assess initial performance to determine substitution, rejection, or deletion errors. The table shows an ideal 100% speech recognition test results-word spoken was the word recognized. A more detailed discussion about factors affecting recognition performance can be found in [3].

Table 1. Vocabulary Confusability Matrix(Credit: Author)

		Actual Response				
		Word 1	Word 2	Word 3	Word 4	Word 5
Predicted Response	Word 1	x				
	Word 2		x			
	Word 3			x		
	Word 4				x	
	Word 5					x

Figure 10. Voice-Control Development Considerations (Credit: Author)

The remaining recognition factors in Fig. 10 are prime candidates to have HSI domain experts engage in. A good way of assessing where HSI is needed during HSI planning is to develop an HSI Applicability Matrix such as shown in Table 2 for a VC system development effort—Recognition Performance Factor (RPF) and NASA HSI Domains. Those HSI areas marked as Prime effort would be where more HSI focus would occur during development. The support effort would be where the other domain SMEs are needed to ensure all system development efforts for the RPF affecting the human are considered. Of importance is to engage a Human Factors Engineering (HFE) SME at the onset to ensure issues like the Shuttle flight experiments lessons learned SSLL 1-6, 8, 10-12 previously mentioned are mitigated. Of note is that when the Shuttle flight experiments occurred, NASA did not have in place the HSI processes and domains. An HFE SME was assigned but late in the project flow when the flight system was already designed. Early incorporation of the HSI domains should ensure that the VC system provides a high level of trust (reliability, robustness, understandability, explication of intention in the feedback) to the user working on the task [22].

Table 2. Notional HSI to VC Applicability Matrix (Credit: Author)

	NASA HSI Domains					
Recognition Performance Factors	**Habitability**	**HFE**	**Training**	**Maintenance and Sustainment**	**Safety**	**Operations**
Task	S	P	S	S	S	S
Environment	P	P	S	P	S	S
Training	S	S	P	S	S	S
Dialogue		P	S	S	S	S
Vocabulary		P	S	S		
	P=Prime effort, S=Support					

The **Task** RPF requires strong HFE participation in developing the VC system application. HFE would be involved at the onset of the task definition and the associated cognitive steps necessary to achieve the task goal(s). These steps and procedures aid in the development of the task mental model in the execution of the task. Habitability SMEs should provide input to the environment expected—acoustics, temperature, location of equipment, etc. Training SMEs support the development of the dialogue by inputting the training required for the task. If the task entails Maintenance and Sustainment (M&S), SME from that domain should support the dialogue and training effort, including electronic procedures or space systems data retrieval. Safety is involved in the VC system development when critical operations are involved as well as doing hazard analysis that may affect the task. For A&A, Operations SMEs would support human/machine resource allocation, task procedure/timeline development, and overall mission operations during the performance of the VC task.

The **Environment** RPF is affected by the habitability of acoustics and the location of the microphone of the VC system. Habitability SMEs can influence the location of equipment, microphones, and/or materials to lessen background noise and reverb effects. These can help improve recognition performance. For spacesuit applications, the helmet poses many challenges in terms of noise and reverb. HFE SME experts would help in developing early mockups and later in the development of the trainers that the astronauts would use. The same would apply for tasks that involve M&S, particularly ensuring the M&S procedures that are brought up via voice are correct in accordance with the task.

The **Training** RPF is a significant factor in ensuring the astronauts are trained adequately on the VC task. Training personnel would be primary in ensuring that the training and the task are optimized for the astronauts. This may mean having HFE personnel involved as well, particularly when human performance assessments (e.g., usability and workload) analysis is done. Training also assesses the mental models of the task to ensure procedures are optimized. Modifications to the dialogue and/or the recognizer may be required. Habitability supports the environment of training to ensure the astronauts are trained in the environment they will use the system. Evaluation of mockup, simulator, and high-fidelity trainers would engage Training SMEs.

In applications that use two forms of system control, such as VC and tactile switches, the training must focus on both forms of system control (Example: the VCS flight experiment that used both VC and the CCTV switch panel to command the CCTV system). The dual training ensures that the user can easily switch back to the other form of control if speech recognition performance has degraded. This is most important when a legacy system in use for a while and it has been retrofitted with VC as an alternative means of control. Users will revert to the control they are comfortable using if the alternative control method such as VC is not sufficiently accurate or usable [23].

The **Dialogue** RPF is how the user perceives the behavior of the VC system. How well the system behaves during the execution of a task will determine the acceptability of the system. Correcting and recovering from an error takes time away from completing the task while specific procedures in the task may not tolerate having to deal with correcting the error [24]. An important part of the dialogue is that it should represent the mental model of the user on how the system should operate. Feedback is an important aspect of the design, letting the user know the state of the system. If not done properly, poor feedback can result in user frustration and possible human errors. Consistency and length of the feedback also affect the user. This is important during times when there is a misrecognition and what the user perceives as the issue. If the state of the VC system misleads the user, this can result in errors that are difficult to recover, especially when the task requires time-critical steps. HFE along with support from the other domains ensures that the dialogue along with the hardware is optimized for the task.

Finally, the **Vocabulary** RPF is developed in conjunction with the task analysis. The application task lead (e.g., spacesuits, robotics workstation) and training work with the HFE and the VC team to optimize the vocabulary. In addition, a linguistic expert may be called to help in the vocabulary as word types, such as one that is multi-syllable with few unvoiced sounds, are best for recognition [7].

7. Conclusion

The two space shuttle voice control flight experiments showed the feasibility of using voice control in space. However, voice control of a spacecraft system presents several challenges. The environment is unlike any other industry that uses the technology, particularly use in a spacesuit. Acoustics can be challenging as spacecraft must be designed to factor in the safety of the equipment and crew and may not consider reverb effects. Adding the stress of being far from the Earth in a hostile environment compounds the problem of getting good speech recognition that contributes to highly reliable voice control. Speech recognition technology has improved considerably but still does not compare to human understanding. Nevertheless, a reliable and robust voice control system application for future space missions is possible with a systematic development approach. By applying systems engineering along with HSI engaged early in the program/project, the goal of a VC space system that can be easily operated and adaptive to the user rather than the user having to adapt to the VC space system is possible.

Acknowledgments

The author would like to thank David Juge and Argun Sethi of the Human-Computer Interface Branch for their technical review of the paper. Also, much appreciation from Ms. Cindy Erdin of Jacobs Engineering for providing a technical review of the paper format and content. Finally, the author wishes to thank the Avionics Systems Division and the Human Interface Branch management for their support in writing this paper.

References

[1] The Artemis Plan, National Aeronautics and Space Administration, September 2020. Retrieved from: https://www.nasa.gov/sites/default/files/atoms/files/artemis_plan-20200921.pdf

[2] B. Gore, (July 31, 2021), Risk of Adverse Outcomes Due to Inadequate Human Systems Integration Architecture, Retrieved from: https://humanresearchroadmap.nasa.gov/risks/risk.aspx?i=175

[3] G. Salazar, "Development Considerations for Implementing a Voice-Controlled Spacecraft System," 2019 IEEE International Symposium on Measurement and Control in Robotics (ISMCR), 2019, pp. B3-1-1-B3-1-20, DOI: 10.1109/ISMCR47492.2019.8955694.

[4] National Aeronautics and Sapce Adminstration. (2021), NASA Human Systems Integration (HSI) Handbook, NASA/SP-202100010952.

[5] J. Payette, "Advanced human-computer interface and voice processing applications in space," in Proceedings of the Workshop on Human Language Technology, Association for Computational Linguistics, Plainsboro, NJ, pp. 416–420,1994.

[6] M. Forberg, 2003, Why is Speech Recognition Difficult, the Chalmers University of Technology, Retrieved from: https://www.huffpost.com/entry/why-is-speech-recognition_b_13873978

[7] K. Kozuch, (July 14, 2020), Alexa vs. Google Assistant vs. Siri: Which smart assistant is best? Retrieved from: https://www.tomsguide.com/us/alexa-vs-siri-vs-google,review-4772.html

[8] Author S., Speech Production System, Retrieved from: http://www.authot.com/en/2016/08/04/speech-production-system/, 2016.

[9] K. Paliwal and K. Yao, Robust Speech Recognition Under Noisy Ambient Conditions, Ambient Intelligence, Chapter 6, 2010.

[10] S. Kim, B. Raj, and I. Lane, "Environmental Noise Embedding For Robust Speech Recognition," Carnegie Mellon University, 2016.

[11] T. Yoshioka, A Sehr, M. Delcroix, K. Kinoshita, R. Maas, T. Nakatani, and W. Kellermann, "Making Machines Understand Us in Reverberant Rooms: Robustness against Reverberation for Automatic Speech Recognition," IEEE Signal Processing Magazine, vol. 29, no. 6, pp.114-126, Nov. 2012

[12] C. A. Simpson, “Speech variability effects on recognition accuracy associated with concurrent task performance by pilots,” Ergonomics, vol. 29, no. 11, pp. 1343–1357, 1986.

[13] C. Yu and, J. H. L. Hansen, "A Study of Voice Production Characteristics of Astronaut Speech During Apollo 11 for Speaker Modeling in Space". Journal of the American Statistical Association, 2017.

[14] K. A. LaBel, Radiation Effects on Electronics 101, Retrieved from: https://nepp.nasa.gov/DocUploads/392333B0-7A48-4A

[15] G. Salazar, STS-41 Voice Command System Flight Experiment Report, Retrieved from: https://ntrs.nasa.gov/api/citations/20180001751/downloads/20180001751.pdf

[16] G. Salazar, D. M. Gaston, and D. S. Haynes, DTO-675: Voice Control of the Closed Circuit Television System. Retrieved from: https://ntrs.nasa.gov/api/citations/20180001752/downloads/20180001752.pdf

[17] National Aeronautics and Space Administration. (2020), NASA Procedural Requirements 7123.1C, NASA Systems Engineering Processes, and Requirements. Retrieved from: https://nodis3.gsfc.nasa.gov/npg_img/N_PR_7123_001C_/N_PR_7123_001C_.pdf

[18] National Aeronautics and Space Administration. (2021), NASA Space Flight Program and Project Management Requirements. Retrieved from: https://nodis3.gsfc.nasa.gov/displayDir.cfm?t=NPR&c=7120&s=5E

[19] NASA/TP-2014-218556, Technical Publication: Human Integration Design Processes (HIDP). NASA ISS Program, Lyndon B. Johnson Space Center, Houston TX, September 2014. Retrieved from: https://nasa.gov/sites/default/files/atoms/files/human_integration_design_processes.pdf

[20] National Aeronautics and Space Administration. (2010), Technology Readiness Levels Demystified, Retrieved from: https://www.nasa.gov/topics/aeronautics/features/trl_demystified.html

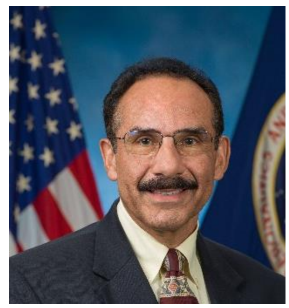
[21] ISO 9241-210- Ergonomics of human-system interaction, Part 201- Human-centered design for interactive systems, 3-15-2010

[22] T. Sheridan, (1988). Trustworthiness of Command and Control Systems. IFAC Proceedings Volumes, Volume 21, Issue 5, Pages 427-431. Retrieved from: https://www.sciencedirect.com/science/article/pii/S1474667017539452

[23] J. Junqua, J., Robust Speech Recognition in Embedded Systems and PC Applications, Springer International Publishing, Chapter 3, 2000.

[24] U. Shrawankar and V. Thakara V., “Adverse Conditions and ASR Techniques for Robust Speech User Interface,” IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 3, September 2011

BIOGRAPHY



George Salazar is the Human-Computer Interface Technical Discipline Lead at NASA, Johnson Space Center, Human Interface Branch focusing on human interfaces for current and future space programs. He attained his bachelor’s in electrical engineering from the University of Houston and his Master of Science in Systems Engineering from Southern Methodist University. He is currently the displays and controls subsystem manager for the commercial crew program and is investigating advanced human-computer interface technology, including graphics processing beyond low earth orbit radiation environment.