Houston Community College

Voice Tech in the Multiverse: Alien Encounter

Group Members: Trevon Woods, Gregory Livingston, Codie Munos, Patrick Pacheco

ITAI 2373 - Module 03

Group 3

Patricia McManus

06/16/2025

Voice Tech in the Multiverse: **Alien Encounter**

Course: ITAI 2373 – Module 03

Group 3 : Trevon Woods, Codie Munos, Patrick Pacheco, Gregory Livingston

Acoustic Challenges & Environmental Factors

- First contact will be on a world similar to the planet Pandora from the movie Avatar.
- New animal sounds.
- Hyper-Thin Atmosphere making it hard for sound to travel without a delay
- Storms causing perpetual thunder, sonic booms from supersonic winds, or a constant "scream" from high-energy particles interacting.

Alien Characteristics

The aliens would look like merfolk and are either skinny or chunky. The skinny aliens are female and have higher pitched voices and the chunkier aliens are male and have deeper pitched voices. They have a vocal/auditory enhancement in how they communicate.

- Communication involves complex, long-range calls or echolocation. This sounds similar to a whale and denotes a sentence being said.
- The sounds change from monotone to sharp depending on the complexity of their “words”.
- They end each sentence with a bleating sound similar to what goats do but much slower; this is to convey emotion or the sentiment of the sentence.

Their limitations could come from being on the surface instead of in the water. This is because sound travels faster and further in aquatic environments so the strength and tone of their vocals would be weaker than usual on land. This could create problems with capturing the correct diction which could lead to incorrect translations.

User Characteristics

- Enhance the hearing frequency of the user to pick up the subtle changes in the alien's speech.
- Enhance hydrophones, sonars, and pressure wave sensors changing the soundwave into human audible translated speech.
- Use device to adapt and learn linguistic priorities for processing sound with different meanings or emphasis such as pitch, rhythm, or tone.
- Recognize changes in alien color or mannerisms.
- Pick up pheromone signals

Noise Sources

Abiotic (Non-Living):

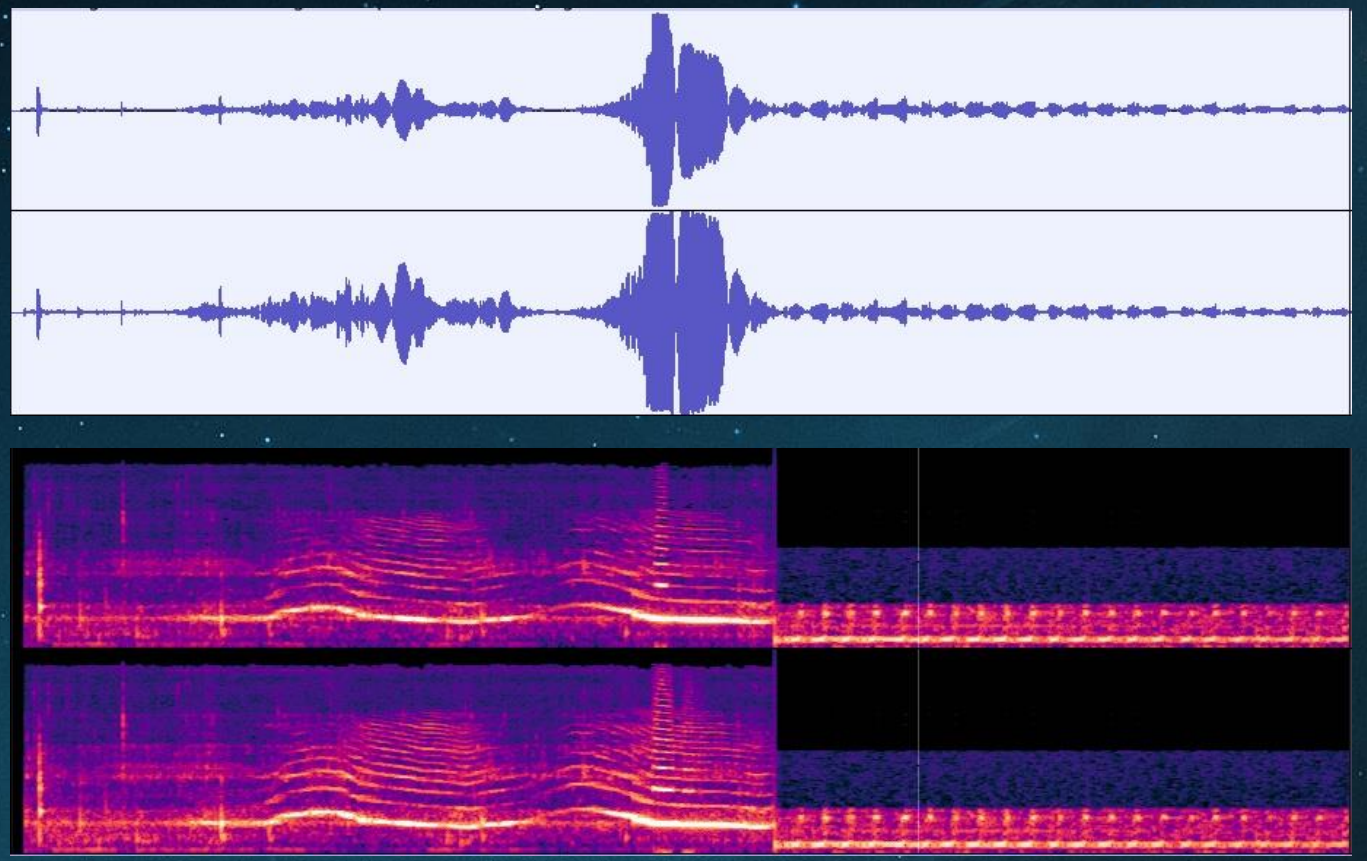
- Howling from the wind
- Sound of the waves
- Weather
- Geysers
- Thermal Vents
- Seismic Activity
- Erosion sounds

Biotic (Living):

- Lifeform calls, clicks, songs, growls, chirps, screeches, ultrasound squeaks, infrasound rumbling, multifrequency
- Lifeform Movements
- Trees and unique foliage

Acoustic Environment Mapping

- Remote sensors
- Probes
- Sonar
- Vibration sensors
- Atmospheric composition analyser
- Landers/Rovers



Non-Human Vocal Anatomy

The aliens we have encountered have gills that rub together like violin strings to produce frequencies out of their mouths for communication.

Universal Translator – Preprocessing Pipeline

Preprocessing Pipeline Flowchart

Why Standard Features Like MFCCs Aren’t Enough?

Traditional audio features like MFCCs are designed for human speech, which typically follows a narrow frequency range and consistent vocal patterns. These assumptions may not apply to alien communication.

Challenges with Alien Speech:

- May include high-frequency or ultrasonic elements
- Could rely on pulses, clicks, or unfamiliar acoustic structures
- Emotional or semantic content might be encoded in pitch curves or vibration timing

Solution:

- Use deep neural networks to learn features directly from raw audio
- Apply spectrogram-based CNNs and self-supervised models (like wav2vec 2.0)
- Let the model learn what matters from alien speech instead of assuming human traits

NLP System Adaptations for Alien-Human Communication

Feature Strategy and ASR/TTS Adjustments

Feature Strategy

- Use transformer models that analyze pitch, duration, and signal shape
- Integrate prosodic features and possible gesture/symbol data

Acoustic Modeling

- Develop a hybrid acoustic model using alien-human paired data
- Use GANs to simulate alien speech patterns for training

TTS (Text-to-Speech)

- Voice generation using neural TTS that mimics tone, emotion, and intonation patterns from alien languages

ASR (Speech Recognition)

- Use sequence-to-sequence NLP models (BERT or Whisper) fine-tuned on alien language

Echo Bridge-

"Breaking barriers, bridging worlds"

Echo Bridge is a gamechanger for interstellar communication! With the tools of text-to-speech, language detection, and AI powered neural networks, our product bridges the gap between alien and human communication! This is perfect for people who are into alien communication as a hobby along with large space agencies who want to go beyond the stars in terms of communication!

Key features:

- Echo Bridge is the first universal translator to be used outside of Earth's realm!
- Our product can decode multiple forms of communications along with use neural networks to continue to evolve!
- We have safeguards in place to protect both humans and alien life!

Competitive Advantages:

- Unlike Earth based translators, Echo Bridge dives deeper into new methods of translations. It is not solely based on words, rather it can retrieve sounds, movement, and patterns to give the best word translation to humans!
- Echo Bridge functions offline rather than needing to be connected to WiFi.
- Our system trains itself which makes it suitable for new methods of communication to be understood.

Demo Scenario

Scene 1: "First Contact"

Narrator:
The crew activates the Universal Translator to speak with a peaceful alien race orbiting Saturn's moon.

Greg:
"Let's open the channel. Translator is live."

Alien:
Strange harmonic tones begin to play.

Scene 2: "Interference"

Trevon:
"Hold up… there's another signal crashing in."

Narrator:
Suddenly, overlapping voices echo. It's the Erasers — the enemy species jamming the line.

Echo Bridges System Output:
"Eeeekk—peace—destroy—friend—run—!"

Scene 3: "Chaos"

Codie:
“The translator’s freaking out. It can’t separate the voices!”

Alien:
Frustrated tones rise. Confusion intensifies.

Patrick:
“They’re using the mimic jammer again!”

Scene 4: "The Fix"

Greg:
“Deploy the fix. Load the Noisy Dual-Speech model.”

Narrator:
The team had trained the model with overlapping alien and human audio — teaching it to lock onto the clean human voice.

Scene 5: "Recalibration"

Translator:
“Processing interference... recalibrating filters...”

Technical Note:
The model overlays mimic noise during training and learns to prioritize the clean signal.

Scene 6: "Clarity Restored"

Translator:
“Message: ‘The Erasers are planning to intercept your next signal.’”

Alien:
Relieved harmonic hums return.

Trevon:
“We’re back online!”

Scene 7: "Warning Delivered"

Codie:
“Sending encoded alert to base. They need backup.”

Narrator:
The translator now ignores Eraser interference completely.

Scene 8: "Victory"

Patrick:
“Too bad, Erasers. Tech’s stronger now.”

Greg:
“Communication saved. Let’s finish the mission.”

Narrator:
With the mimic jammer neutralized and the translator stronger than ever, humanity's alliance with the peaceful alien race continues. The Erasers retreat into static silence—for now.

BONUS: Villain’s Countermeasure

"The Mimic Jammer"

A device that imitates human and alien voices at once, confusing the translator’s speech recognition system.

It scrambles frequencies with overlapping emotional tones, making it difficult for the Universal Translator to determine who is speaking or what is being said.

"The Erasers"

The Mimic Jammer was created by a group of hostile alien species known as the Erasers.

The Erasers despise the alliance between humanity and these peaceful extraterrestrials, and they seek to disrupt all forms of communication between us.

Solution:

We used a method called Noisy Dual-Speech Training, where we deliberately fed the model confusing audio. The translator learned not to get overwhelmed when it hears two voices at once. Here's how the training process worked:

- **Generate** clean human speech samples
- **Overlay** alien noise or mimic voice on top
- **Label** the correct (human-only) transcription
- **Train** the model to recognize and prioritize the clean voice, ignoring the interference

"No Jam Can Hold Us" Rap about Fixing the Translator

Bonus Opportunity

Intro – Patrick Talking

They tried to block the peace. Corrupt the code.
But we trained the machine…To hear the truth in the noise

Hook- Gregory Livingston

They jammed the line ,they some fakers
Stealin' voices, like imitators
We rebuilt the code, now it understands
The mimic jammer was a stupid plan
The Erasers took a L, and need a new plan…
made them back off — and stop playin

Verse1 Trevon Woods

Bad aliens lurkin, Erasers ready to murk em.

They mimic our voice, because they had no choice.

Translator broken, can't get these tokens.

Still, we don't fold, we make our stand bold.

We built a system, trained it deep, they can't stop us we don't know defeat.

Verse 2 Codie

We fed the noise to our translator,
Trained our model about the fakers
The pretenders speak, but their meaning in danger.
Locked and bound to the warping of sound.

Trained the model on what to take
Gave it the edge that no jam could break.
Now when two signals clash and fuse,
The erasers will crumble in the feud