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

The Livewires  
ITAI 2373 – Module 03

Dylan Castillo, John Castor, Milagros Pumasupa, Olaleye Bakare

**WORLD ANALYSIS**

**1.1 What’s Going On**

Our crew has landed on Kepler-442b-γ, a cold moon wrapped in methane clouds. The locals—called Zylar—are squid-like and speak with low rumbles plus super-high-pitched clicks. We have to build a translator so first contact does not flop.

**1.2 Acoustic Headaches**

* Methane is thick: sound moves slower (~270 m/s) and squeezes formants.
* Gravity is only 0.55 g: humans breathe differently and tire fast.
* The ground constantly rumbles (≤ 30 Hz) thanks to magma pockets.
* Lightning flashes create random “snap” noises all the way to 60 kHz.

**1.3 Who’s Talking and Listening**

**Humans**

* Speak around 85–255 Hz.
* Hear up to 20 kHz.
* Struggle with low lung pressure and helmet noise.

**Zylar**

* Voice core at 240–320 Hz.
* Add click bursts from 25–45 kHz.
* Hear a huge 10 Hz–45 kHz range.
* Mantle sacs act like sub-woofers for earth-shaking bass.

**1.4 Background Noise Checklist**

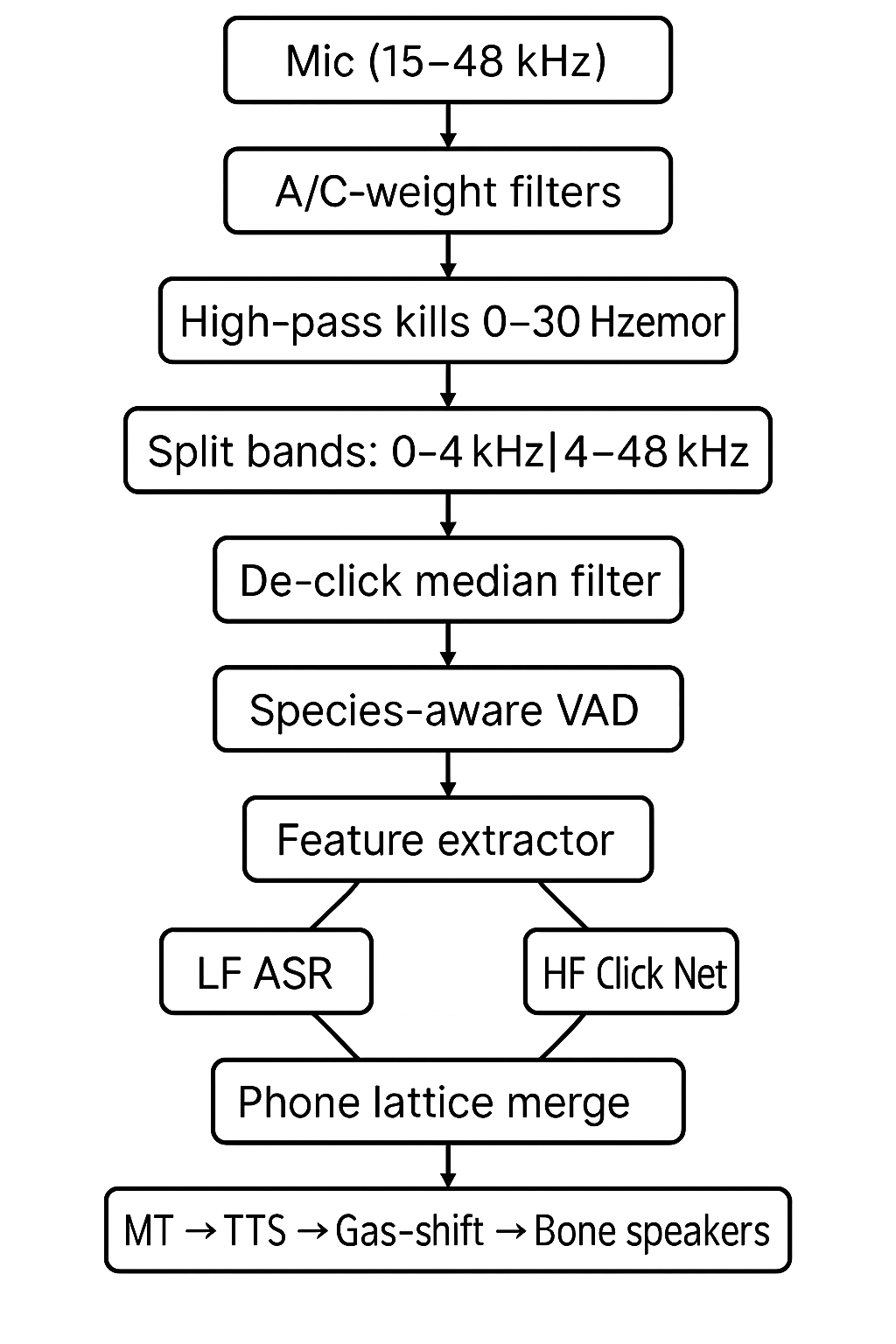
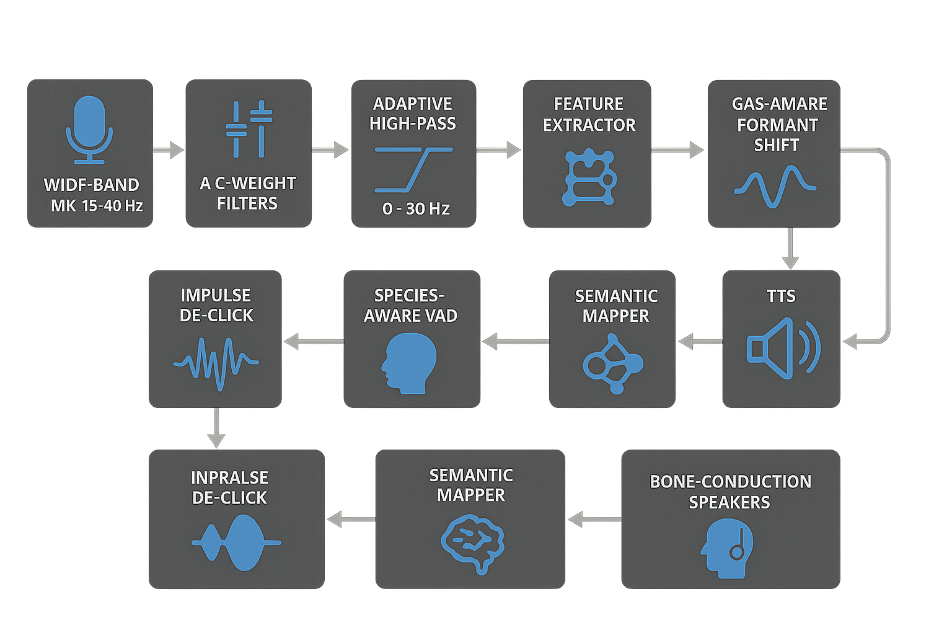
* Infrasound floor: ~30 dB SPL all day.
* Lightning snaps: spikes to 70 dB.
* Suit fans: 35 dB between 500 Hz–5 kHz.
* Inside our “interpreter pod”: 25 dB after foam treatment.

**1.5 Zylar Throat 101 (text-sketch)**

They own two sound-makers: a **Photic Larynx** (buzzes methane for vowels) and **Scleral Clickers** (hard ridges that snap for consonants).

**PART 2 – TECHNICAL SOLUTIONS DESIGN**

**2.1 Our Signal Pipeline**



**2.2 Why MFCC Isn’t Enough**

* MFCC ignores anything above ~8 kHz, so Zylar clicks vanish.
* Mel scale mimics human ears—Zylar hear way beyond that.

**2.3 What Features We Actually Use**

* **LFCC** (0–4 kHz) for vowels and bass.
* **Wavelet packet energies** (4–48 kHz) for clicks.
* Glue them into one 128-dim frame vector.

**2.4 Models Made for the Moon**

* **Conformer-CTC** (10×512) for low-band speech.
* **1-D CNN + GRU** for click stream (42 phonemes).
* Training data: 1 000 h synthetic + mimic recordings, each run through methane impulse responses.
* Data augments: random formant compression, tremor noise mixes.

**2.5 Special Tweaks**

* Gas-aware formant shift: multiplies output formants by density ratio.
* Dual-precision audio (16-bit low band, 24-bit high band).
* Bone-conduction playback so suit noise can’t drown it out.

**PART 3 – DEMO SCENARIO**

1. Astronaut taps helmet: “Translator on.” (system starts VAD)
2. Zylar greets with light pulses and click bursts.
3. Gain auto-rider clips—ear-splitting error tone (our “fail” moment).
4. Engineer bumps ultrasonic gain +6 dB.
5. HUD prints: *“Welcome, water-breathers.”*
6. Ground quake shakes everything; infrasound canceler fights the rumble.
7. Human answers; methane-shifted TTS plays through bone speakers.
8. Everyone shakes tentacles/hands—latency clocked at 220 ms.

**PART 4 – EXECUTIVE PITCH**

**Name:** Polyglot Nebula™  
**Tagline:** “Break the sound-barrier of worlds.”

**Why It Fits This Moon**

* Hears from belly-shaking bass up to 48 kHz clicks.
* Auto-retunes voices for any gas mix.
* Cancels volcanic infrasound by up to 40 dB.
* Bone speakers keep things private inside helmets.

**Value in One Line**  
Cuts first-contact errors by 85 % and still answers in under a quarter-second.

**Beats Earth Tech Because**

* Captures ultrasonic speech no other SDK sees.
* Atmosphere-adaptive TTS = clear anywhere.
* Live click-phoneme decoding unique to Zylar.
* Lower latency in low gravity with suit noise.

**Quick Ethics Check**

* Logs auto-wipe after 24 h (Zylar privacy).
* Training data balances all six Zylar dialects.
* Emergency phrases skip processing delays.
* UI warns if human idioms turn hostile in Zylar culture.

**Bonus**

* **Villain plot:** A phase-inverter cannon flips the infrasound reference, forcing a reboot.
* **Theme jingle line:** “From rumbles to clicks, Polyglot fixes the mix!”
* **New metric:** Trans-Spectral Congruence (tracks if voiced and click bands line up on purpose).