

# VoiceScape Quest

## Initial Concept Submission

---

### Brief Summary

**VoiceScape Quest** is an immersive, voice-driven AR music experience built for Snap Spectacles. Users create synthwave-inspired music loops through voice commands, while neon fractals and visual effects dynamically appear in their real-world space. This hands-free interaction allows for real-time audio visualization, making music production an intuitive and accessible experience.

- **Core Innovation:** A hands-free voice-controlled music sequencer in AR.
- **Key Features:** FFT-based real-time audio visualization, voice input interaction, and procedural fractal animation.
- **Session Length:** ~10–15 minutes per play, designed for repeat engagement.
- **Spectacles Integration:** Uses voice commands, world mesh tracking, and dynamic AR overlays for an immersive, spatially-aware experience.
- **Retention Strategy:** Unlockable music loops, fractal styles, and shareable clips keep users engaged.

### Why This Matters

This Lens redefines how users create and visualize music in AR—merging interactive sound composition with AI-powered fractal animations. With Spectacles' hands-free capabilities, we push AR beyond passive filters into an expressive, creative tool that adapts to user-generated content.

---

### Team Information

- **Michael Chaves** – Lead Voice Interaction & UX Designer, Developer
  - **Sinchana Nama** – Lens Studio Developer/Scripting Specialist
  - **Christopher Park** – 3D Asset Creator (Modeling & Basic Animation)
  - **Paul Reitz** – Backend Development/Optimization
-

# 1. Product Vision

VoiceScape Quest is an immersive, voice-driven AR music experience designed for Snap Spectacles. Users generate synthwave-inspired music loops while audio-reactive fractals and neon visuals dynamically appear in their real-world space.

## How It Feels to the User:

- **Empowering & Accessible** – Users interact entirely through voice commands, making it easy to engage without physical controllers.
- **Immersive & Expressive** – Visual fractals and neon trails grow in sync with the user's music, creating a responsive digital world.
- **Short & Replayable** – Each session is designed for bite-sized, creative interactions (~10–15 minutes), encouraging multiple plays.

**Outcome:** Users leave with a personalized AR soundscape, eager to return and refine their music-driven environment.

---

# 2. Target Audience & User Motivation

VoiceScape Quest is built for:

- **Casual Creators & Music Enthusiasts** – People who enjoy playful experimentation with music loops and AR visuals.
- **Early Spectacles Adopters & XR Enthusiasts** – Users familiar with AR interactions, seeking fresh, engaging ways to interact with their space.
- **Accessibility-Focused Users** – Individuals who benefit from hands-free, voice-driven interactions, making AR more inclusive.

## User Priorities:

- ✓ **Ease of Interaction** – Intuitive voice commands streamline the experience.
- ✓ **Engagement & Personalization** – Users create unique sound-visual combinations each session.
- ✓ **Replayability & Social Sharing** – Users return to experiment and share their creations.

## Similar AR Experiences & Lens Examples

1. **"Electronauts" (VR - Oculus)** – Allows users to create music in a futuristic environment.
  - **Exemplary Feature:** Intuitive music interaction that feels natural in an immersive space.

- **Link:** <https://survios.com/electronauts>
  - 2. **"My Mandala Lens" (Snap AR)** – Uses real-time hand triggers to modify visuals.
    - **Exemplary Feature:** Hands-on interaction creating dynamic visuals.
    - **Link:**  
[https://www.snapchat.com/lens/930cf584ea5b4224803a9f6a339595a6?sender\\_web\\_id=56adc84a-702c-406f-8b0c-643ce1d6d820&device\\_type=desktop&is\\_copy\\_url=true](https://www.snapchat.com/lens/930cf584ea5b4224803a9f6a339595a6?sender_web_id=56adc84a-702c-406f-8b0c-643ce1d6d820&device_type=desktop&is_copy_url=true)
  - 3. **"WZRD: AI Sound Visualizer" (Mobile AR)** – Uses audio analysis to generate reactive light and shapes in AR.
    - **Exemplary Feature:** Real-time music-to-visual pipeline implementation.
    - **Link:** <https://wzrd.ai/>
- 

## 3. Core Mechanics & Interaction Model

### A. Voice-Driven AR Interaction

Users create music and trigger visual changes using simple, natural voice commands:

- 🎵 **"Add Bassline"** – Introduces a synth bass layer.
- 🥁 **"Change Drums"** – Replaces existing drum loops with another variation.
- 🎨 **"Grow Fractals"** – Expands the neon fractal designs in the AR space.
- 🌈 **"Change Colors"** – Shifts the environment's color palette dynamically.

### B. Real-Time Audio FFT & Visualization

- **Music-Driven Graphics:** Uses Fast Fourier Transform (FFT) analysis to analyze the user's track and generate neon fractals and waveforms.
- **Fractal Growth System:** Expands visual patterns based on volume, beat intensity, and frequency range.

### C. AR Principles & Spectacles Optimization

- ✓ **Uninhibited:** Users can freely move around their space, interacting without controllers.
  - ✓ **Interactive:** The Lens dynamically responds to music and voice commands.
  - ✓ **Spatial:** Uses world mesh tracking to anchor visuals realistically to floors and walls.
- 

## 4. User Journey & Flow

### Step 1: Lens Launch & Tutorial (0–1 min)

- A quick on-screen guide explains basic voice commands and how to interact.
- Users can choose a starter synth loop to begin creating.

## Step 2: Music Creation & Visual Expansion (2–10 min)

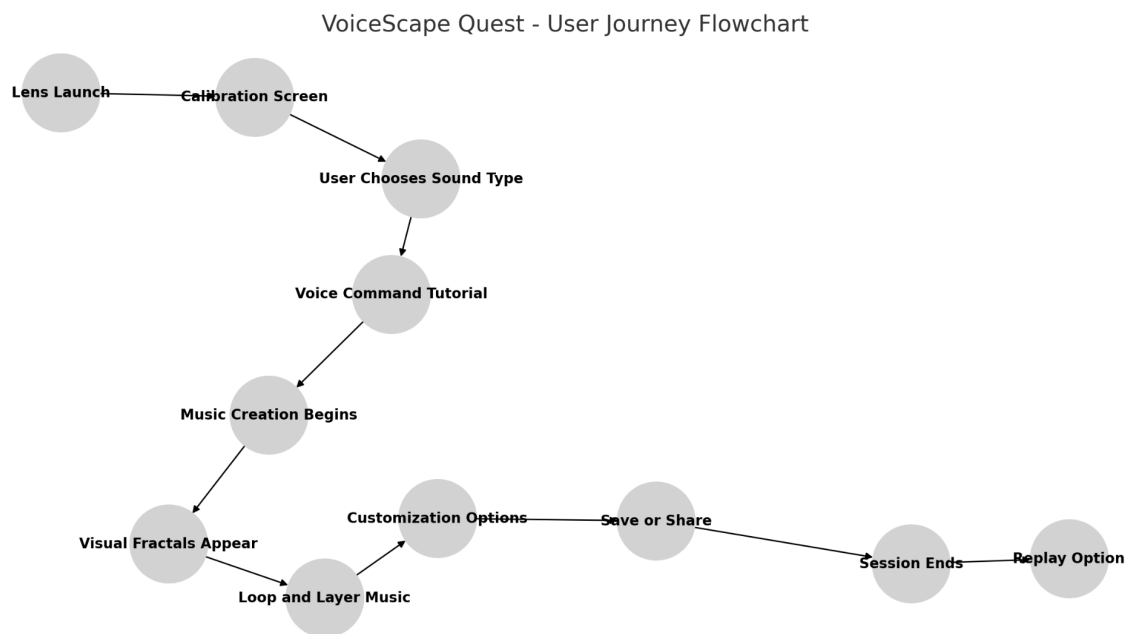
- Users layer new tracks, change loops, and modify visuals using voice.
- Audio-reactive fractals and neon lights expand dynamically in the user's space.

## Step 3: Session Wrap-Up & Sharing (10–12 min)

- Users see a snapshot of their final visual creation.
- Option to save a short clip or share with friends via Snap's social features.

## Step 4: Replay & Retention Hooks (12–15 min)

- Users unlock new color themes, sound loops, or fractal effects by returning for future sessions.



## 5. Retention & Engagement Strategy

To encourage repeat usage, the Lens includes:

- ✓ **Progression & Collectibles** – Unlock new fractal styles or music presets over time.

- ✓ **Surprise Elements** – Small randomized effects per session ensure variety.
  - ✓ **Social Sharing & Snapcodes** – Users can share their AR music sessions and invite friends to try the Lens.
  - ✓ **Replayability Hooks** – The neon fractals and visuals change based on different sound combinations, making each session unique.
- 

## 6. Technical Feasibility & Constraints

### Platform-Specific Adjustments (Spectacles AR Capabilities)

- ✓ **Battery & Thermal Optimization** – Sessions are kept under 15 minutes to prevent overheating.
- ✓ **No Face/Hand Tracking Dependencies** – Voice and world-tracked visuals maximize compatibility.
- ✓ **Efficient Processing** – Visuals are low-poly and use GPU-optimized shaders for real-time fractal growth.

### Development Approach

1. **Lens Studio Audio Processing** – Implement FFT-based analysis using Snap's supported sound interaction tools.
  2. **Performance Optimizations** – Pre-cache lightweight 3D elements for stable frame rates.
  3. **Minimal UI Clutter** – Rely on gesture-based UI prompts instead of traditional menus.
- 

## 7. Timeline & Milestones

### Phase 1: Prototype (February–April)

🎯 **Goal:** Core interaction loop functional in Lens Studio.


- Implement basic voice command framework.
- Prototype real-time fractal growth based on audio input.
- Early internal testing & feedback collection.

### Phase 2: MVP Build (May–June)

🎯 **Goal:** Expand user interaction depth.

- Improve voice recognition accuracy.
- Introduce more music loops & fractal presets.
- Test usability & accessibility for hands-free controls.

### Phase 3: Final Production (July–September)

 **Goal:** Polish, optimize, and prepare for release.

- Finalize graphics, UI animations, and performance tweaks.
  - Implement feedback from Snap’s mentorship sessions.
  - Submit to Snap AR for review & potential publication.
- 

## 8. Budget & Resource Planning

### Estimated Budget Allocation

- **Lens Studio Development: \$3,000** – Implementation of audio FFT, UI flow, and performance optimization.
- **3D Asset Creation: \$2,500** – Custom fractal models, neon elements, and UI assets.
- **User Testing & Iteration: \$1,500** – Accessibility testing and feedback implementation.
- **Miscellaneous (Software & Cloud Services): \$1,000** – Voice processing tools, rendering optimization.

 **Total Estimated Budget: \$8,000**

---

## 9. Success Criteria (HEARTS Framework)

- ✓ **Happiness** – Users feel engaged and excited by the music-visual integration.
  - ✓ **Engagement** – Voice interaction makes it easy to experiment with soundscapes.
  - ✓ **AR Intent** – Only possible in AR wearables, making real-world spaces part of the music.
  - ✓ **Retention** – Short, replayable sessions encourage multiple visits.
  - ✓ **Task Success** – Commands are intuitive and responsive.
  - ✓ **Spectacles Vision** – The experience is optimized specifically for Spectacles' hands-free interface.
-

## 10. Spectacles Feature Integration

- ✓ **Voice-Driven UI:** Enables a fully hands-free experience.
- ✓ **Audio Visualization:** FFT-based sound processing drives neon fractals and lights.
- ✓ **World Mesh Anchoring:** Ensures AR elements feel integrated into real-world space.

By aligning VoiceScape Quest with Snap's AR principles, retention strategies, and technical capabilities, we're creating a unique, voice-driven AR music experience that thrives in the Spectacles ecosystem.