SpeakEasy

Voice-Driven AI for Inclusive XR

Michael Chaves
M.Des Thesis Process Book
San Jose State University

for Mason

San Jose State University
M.DES EXPERIENCE DESIGN
MASTER THESIS PROJECT

SpeakEasy

Voice-Driven AI for Inclusive XR

designed by
MICHAEL CHAVES

advised by
YOON CHUNG HAN, Ph.D.
Associate Professor, SJSU

MARJAN KHATIBI
Assistant Professor, SJSU

DYLAN FOX
Director of Operations, XR Access

TABLE OF CONTENTS

INTRODUCTION

ABSTRACT

Extended-reality headsets promise worlds that bend to human imagination, yet they still ask many users to grip two plastic controllers and master an invisible grammar of buttons. For individuals with low muscle tone, that first request can shut the door on the entire medium. **SpeakEasy: Voice-Driven AI for Inclusive XR** asks a simple question: ***What if spoken language, our most familiar interface, became the primary way to act inside mixed reality?***

The project unfolded through a research-through-design journey that combined literature review, eight participatory co-design sessions, Wizard-of-Oz voice trials, and three full-scale mixed-reality prototypes. Each cycle refined a framework that is now presented as the **Ten Pillars for Accessible Voice Interaction in Spatial Computing**. This framework considers everything from microphone placement and intent confidence thresholds to empathetic feedback and cultural sensitivity.

Building on these insights, the final prototype delivers three guided wellness modules that respond to natural language. This demonstrates that accessibility can be achieved with readily available software rather than requiring custom hardware.

Evaluation with eight target users produced a 79 percent task-completion rate and a score of 79 out of 100 on a standardized usability questionnaire (System Usability Scale, or SUS). These results indicate above-average learnability and satisfaction. Qualitative feedback revealed moments of delight when participants realized they could explore and manipulate space without lifting a finger.

These findings suggest that natural speech, combined with real-time AI personalization, can provide inclusive and embodied immersion without compromising craft or engagement. SpeakEasy positions voice as a viable primary input for spatial computing and offers designers a practical roadmap toward multimodal XR experiences that welcome more bodies, voices, and stories.

LOCATING THE THESIS RANGE

Design for immersive technologies still privileges users who can operate two hand-held controllers. The 2019 American Community Survey reports that 19.9 million U.S. adults have limited grip strength or other upper-limb mobility challenges, making conventional XR input exclusionary from the very first click (U.S. Census Bureau, 2019). Recognizing this gap, this thesis explores how voice and other alternative input methods can create more accessible and engaging XR experiences.

PRIMARY SUBJECTS

Here are three primary subjects that frame this investigation:

- Accessibility in Technology – reducing physical and cognitive barriers.
- Extended Reality (XR) – expanding interaction beyond screens.
- Inclusive Design – centering diverse bodies and voices from the outset.

QUALITIES THAT LINK THESE INTERESTS

- Formal: Voice and spatial computing shape the interaction layer.
- Historical: Accessibility guidelines have evolved from ramps to responsive UI to voice.
- Conceptual: Human dignity and autonomy sit at the core of experience design decisions.
- Intuitive: Conversational input mimics real-world social exchange.

ASSOCIATED TOPICS FOR EVIDENCE GATHERING

- Universal Design standards in digital ecosystems.
- Market penetration of voice assistants and headset adoption.
- Therapeutic outcomes of voice-guided mindfulness applications.

ASSOCIATED TOPICS FOR EVIDENCE GATHERING

- XR can transcend traditional access barriers by removing the need for fine-motor input.
- Real-time AI personalization adapts experiences to individual speech patterns and intent confidence.

U.S. Census Bureau. (2019). *American Community Survey 1-year estimates, Table C18108: Disability characteristics by ambulatory difficulty* https://data.census.gov

LOCATING THE THESIS RANGE

REFLECTION ON FEEDBACK

Committee input stressed that accessibility is not just functional but emotional. Early prototypes focused on task completion rates; the final iteration asks how voice interaction can foster calm, agency, and joy for users with limited mobility. This pivot led to participatory sessions with eight target users, where qualitative insights reshaped the tone, pacing, and error handling of the SpeakEasy dialogue model.

UNEXPLORED THREAD NOW IN FOCUS

How does a voice-first interface influence a user's sense of autonomy and social presence in shared mixed-reality spaces? Future work will adopt diary studies and longitudinal testing to capture these psychological effects.

REVISED PROJECT STANCE

SpeakEasy positions voice as a primary channel of agency rather than a convenience layer. By coupling natural-language understanding with real-time scene adaptation, the prototype offers an accessible, intuitive, and engaging path to mixed-reality interaction for a broader spectrum of users.

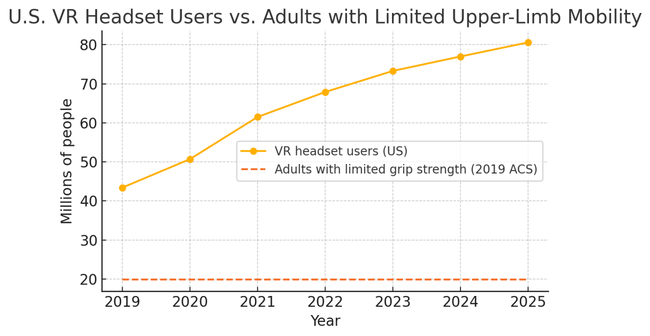


Figure 1.1 — U.S. XR Accessibility Gap
Line chart comparing VR-headset adoption (2019 – 2025) with adults who report limited upper-limb mobility (19.9 million). Source: DemandSage.

DemandSage. (2024). *Virtual reality statistics: Users and adoption forecasts 2019–2028* (compiled from Oberlo/Statista projections) https://www.demandsage.com/virtual-reality-statistics

THESIS STATEMENT

Natural speech, when paired with real-time AI personalization and accessibility-first design principles, can serve as a primary mode of interaction in mixed reality. SpeakEasy demonstrates that users with limited upper-limb mobility can confidently navigate spatial computing experiences through voice input alone.
This approach provides accessible, engaging, and immersive interaction, and points toward a future where multimodal XR experiences are inclusive of more bodies, voices, and perspectives.

PEOPLE, PLACES, AND THINGS

3 PHYSICAL PLACES

- Accessibility Labs: Research and development spaces focused on creating accessible technology solutions.
- Universities: Academic institutions with departments dedicated to inclusive design and technology.
- Tech Conferences: Events where the latest innovations in XR and accessibility are showcased.

3 HUGE OBJECTS

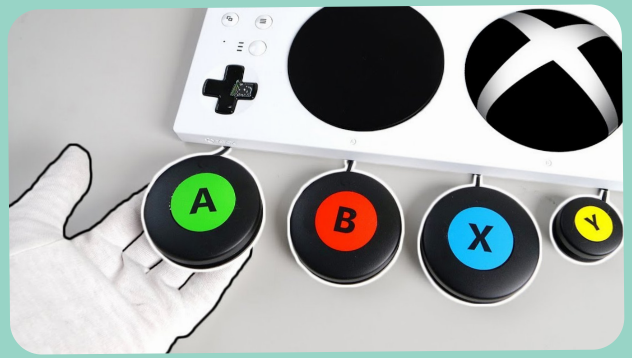
- Virtual Reality Headsets: Key tools in XR experiences, needing adaptation for accessibility.
- Adaptive Controllers: Devices designed to make gaming and interaction within XR environments accessible.
- Gesture Recognition Systems: Technology that interprets human gestures as commands, crucial for accessible design.

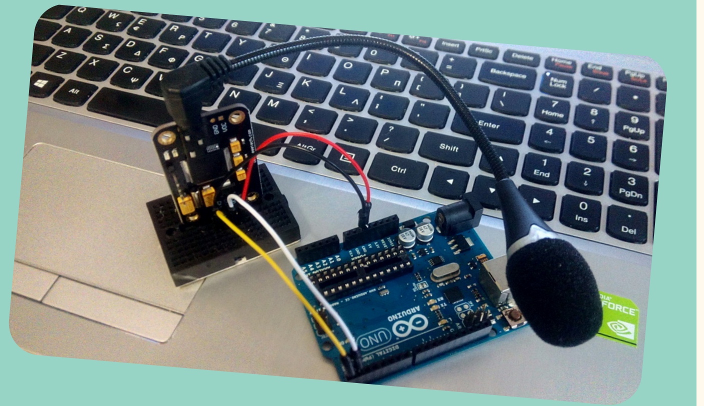
3 TINY OBJECTS

- Braille Displays: Devices that allow visually impaired users to read text through tactile feedback.
- Haptic Feedback Devices: Small wearable gadgets that provide sensory feedback to enhance XR experiences for users with hearing impairments.
- Voice Command Modules: Compact devices that interpret voice commands, enabling hands-free interaction within XR environments.

3 ABSTRACT QUALITIES

- Inclusivity: Ensuring XR environments are welcoming and usable for all.
- Empowerment: Giving users control over their XR experiences.
- Adaptability: The ability of XR technologies to cater to a wide range of needs and preferences.





3 PHYSICAL QUALITIES

- Accessibility Features: Built-in options that make XR tools usable by people with disabilities.
- User Interface Design: The layout and design of XR applications that accommodate diverse users.
- Sensory Adaptation: Modifications in XR that allow for multi-sensory engagement.

3 FOUNDATIONS OR ORGANIZATIONS

- XR Access Initiative: Works to make XR technologies accessible.
- Global Initiative for Inclusive ICTs (G3ict): Promotes inclusive ICTs worldwide.
- World Wide Web Consortium (W3C): Develops international standards for web accessibility, including XR.

3 HISTORICAL EVENTS

- The Adoption of the Web Content Accessibility Guidelines (WCAG): Set standards for digital accessibility.
- The Launch of the First VR Headset: Marked the beginning of the modern XR era.
- The Enactment of the Americans with Disabilities Act (ADA): Major legislation promoting accessibility in all spheres of public life.

3 SUPERMARKET ITEMS

- Smart Glasses: Augmented reality glasses that could be adapted for accessibility in everyday tasks.
- Voice-Activated Assistants: Products that facilitate voice control, making technology more accessible.
- Wearable Fitness Trackers: Devices that could be integrated into XR for health monitoring and interactive experiences.

RESEARCH

RESEARCH QUESTIONS

My journey into voice-driven XR began with a simple but urgent inquiry: ***how can we design mixed reality experiences that invite more users in, rather than keeping them out?*** As an experience designer and creative technologist, I set out to explore how conversational interfaces could lower barriers, support agency, and foster a greater sense of belonging for people often excluded by controller-based systems. This study investigates not only what makes voice interaction accessible, but also how it can shape more meaningful, intuitive, and satisfying immersive experiences.

Based on these aims, the following research questions guided this work:

How can voice-driven interaction reduce physical and cognitive barriers in mixed reality for users with low muscle tone?

What design strategies most effectively support inclusive and intuitive voice interfaces in XR environments?

How does real-time AI personalization influence user engagement, comfort, and agency during voice-guided XR experiences?

In what ways do users describe their sense of immersion, autonomy, and satisfaction when interacting with XR prototypes through voice compared to traditional input methods?

KEY POINTS

Accessibility in XR must be addressed collaboratively, not in isolation.

True inclusion requires narrowing scope and working directly with target communities.

Voice and AI offer real promise, but only when tuned to real user needs and contexts.

The most impactful progress comes from iterative, user-driven experimentation.

FRAMING REFERENCE

SPRING 2024 - THE EXPERT CONVERSATION

To deepen my understanding of how voice-driven interfaces might expand accessibility in XR, I conducted an in-depth interview with **Dylan Fox**, an advocate, technologist, and co-founder at XR Access. Dylan's dual experience as both designer and community builder provided a unique window into the challenges and opportunities of accessible spatial computing. My aim was to ground the project's direction in lived expertise rather than abstract theory.

FRAMING REFERENCE

SPRING 2024 - THE EXPERT CONVERSATION

THE QUESTIONS

- How are accessibility needs currently prioritized in XR design and development?
- Where are the most significant gaps for users with motor or cognitive differences?
- What roles do voice and AI have in bridging these barriers?
- How do practitioners balance idealism with the practical constraints of product development?

The interview was a conversation rather than a Q&A—part story exchange, part design critique. I shared my personal motivations, such as my son's and my partner's children's lived experiences with disability and neurodivergence, and described my journey from game development to inclusive experience design.

Dylan reflected on the sheer breadth of accessibility challenges in XR, noting that even organizations dedicated to the cause must "pick and focus" rather than solve everything at once. He stressed the value of partnering with user advocacy groups and disabled communities directly, cautioning against designing in isolation:

> **"You see all the gaps and you want to plug all of them but yeah, it's challenging at times to pick and focus on things. I think that's one of the things I've been realizing doing this at XR Access. It's like we are trying to kind of find all those things at once and just like connect to the different people that can solve them."**

Another key theme was the tension between technological ambition and realistic scope. Dylan advised beginning with targeted, meaningful prototypes and making room for community-driven iteration. He shared examples of XR initiatives that paired design teams with real users—including quadriplegic and neurodiverse participants—and underscored the importance of co-design:

> **"Always in a case like this trying to start from a community of actual disabled users, and work with them and you know, have them be part of that process and shaping the product is always going to be better results than trying to just kind of brainstorm on your own."**

This conversation directly influenced my research methods, leading me to prioritize participatory design, expert engagement, and real-world user feedback throughout the project. Insights from Dylan and the broader XR Access community were instrumental in defining not only the technical aims of SpeakEasy, but also the core values of agency, co-creation, and empathy that shape its final framework.

SURVEY & RESULTS

To better understand accessibility needs and user attitudes toward voice-driven interfaces in XR, I conducted a survey between February and April 2024. The survey was distributed online, reaching respondents through LinkedIn, Discord, email, and XR Access channels. All responses were collected remotely, with the majority of participants based in the United States, and a few international respondents from Canada and Europe.

A total of 42 respondents participated, representing a cross-section of the XR community:

- Backgrounds included: XR professionals, designers, accessibility advocates, graduate and undergraduate students, and early adopters of immersive technology.
- Experience levels ranged from beginners to advanced users, with many reporting professional or academic involvement in digital media, interaction design, or computer science, and others engaged as hobbyists or through coursework.

The survey used a mix of open-ended reflection and quantitative questions to capture both narrative and measurable insight:

- 28.6% of respondents (12 out of 42) specifically identified voice command as the most useful feature or their top wish for XR applications.
- Approximately 40% reported physical challenges, such as discomfort, grip fatigue, or trouble with manual controllers, highlighting persistent accessibility barriers.
- A similar proportion reported no significant barriers, reflecting a broad spectrum of user experiences.
- On a 1–5 scale, the average rating for how well XR currently meets "unexpressed needs" was 2.8, suggesting there is considerable room for improvement.

Qualitative responses described the desire for more seamless, hands-free, and natural ways to interact—often citing voice as "the most natural" or "the most inclusive" approach. This feedback directly influenced the design priorities for SpeakEasy, reinforcing the need for intuitive voice commands and reducing dependence on manual input.

QUESTIONS

Describe your first XR experience. What stood out?

Recall a moment you felt a strong emotional connection with XR. What triggered it?

Discuss challenges faced due to disabilities in XR.

How did you adapt to using XR technologies? Were there specific features or tools that helped?

What personalization features do you find most useful or wish were available in XR applications?

Have XR technologies impacted your daily routines or behaviors? If yes, please describe how.

What improvements would make XR technologies more accessible and enjoyable for you?

On a scale from 1 to 5, rate how well XR meets unexpressed needs

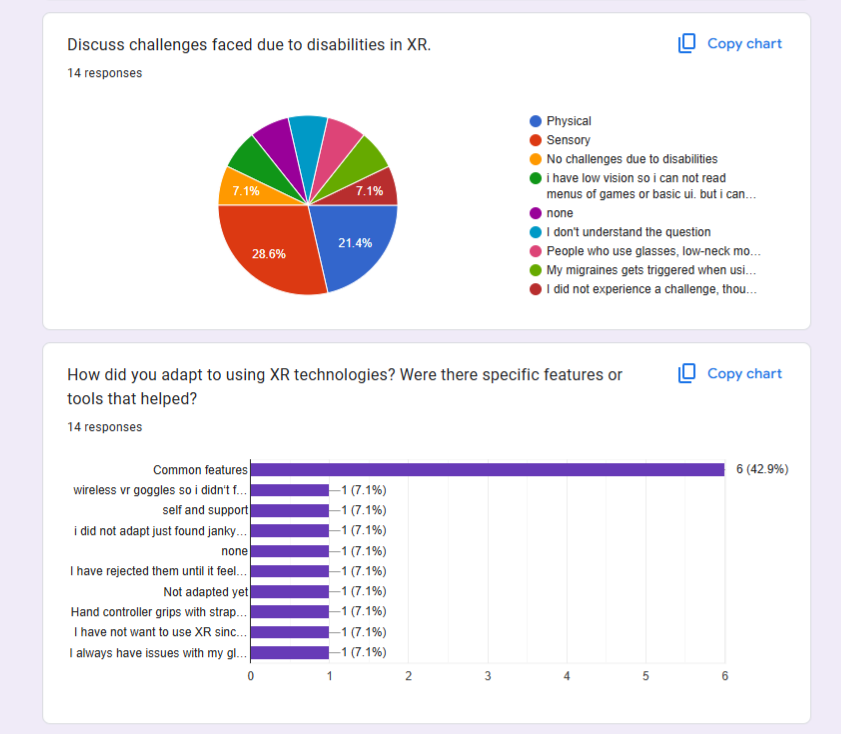


Figure 2.2 — Survey Dashboard
A bar chart summarizes the proportion of users wishing for voice-driven features versus those reporting discomfort or difficulty with controllers, highlighting the gap between current XR solutions and user needs.

REPOSITORY OF INSIGHT

My investigation into the integration of voice-driven AI in XR environments has been informed by a diverse range of perspectives on usability, accessibility, and technology adaptation.

This repository collects influential thoughts and findings that highlight the challenges and potentials of making XR systems more user-friendly and inclusive.

COLLECTION OF QUOTATIONS

“I wish for a world that views disability, mental or physical, not as a hindrance but as unique attributes that can be seen as powerful assets if given the right opportunities”. **- Oliver Sacks**

“We should celebrate neurodiversity — the world would be poorer and life duller if we were all the same”. **- Neil Milliken**

“Let’s stop ‘tolerating’ or ‘accepting’ difference, as if we’re so much better for not being different in the first place. Instead, let’s celebrate difference, because in this world it takes a lot of guts to be different”. **- Kate Bornstein**

“The only disability is when people cannot see human potential”. **- Debra Ruh**

“If disabled people were truly heard, an explosion of knowledge of the human body and psyche would take place”. **- Susan Wendell**

“Design makes us feel empowered. Design makes us independent. Design makes us able, or it doesn’t. And when it doesn’t, it’s not Design. It’s Design failure. It’s not inclusive. It’s exclusive.” **- Pattie Moore**

“The power of the Web is in its universality. Access by everyone regardless of disability is an essential aspect”. **- Tim Berners-Lee**

“It is not enough that we build products that function, that are understandable and usable, we also need to build products that bring joy and excitement, pleasure and fun, and, yes, beauty to people’s lives.”- **Don Norman**

SELECTED QUOATION AND EXPLANATION

The quotation from **Tim Berners-Lee**, “The power of the Web is in its universality. Access by everyone regardless of disability is an essential aspect,” profoundly resonates with the essence of my thesis on enhancing XR accessibility.

By advocating for a universally accessible web, he speaks to the core of inclusive design—creating XR platforms that everyone, including those with disabilities, can access.

This approach views diversity as a catalyst for innovation and insists on empathy in design to ensure dignity and participation for all users.

ANNOTATED BIBLIOGRAPHY ENTRIES

Given the format and instructions, I will incorporate these quotes into the annotated bibliography, detailing the source, context, and relevance of each quote to the broader theme of XR accessibility. This process will involve analyzing the quotes within the framework of existing research, design practices, and technological developments related to inclusivity in digital environments.

MANIFESTO

Design in XR extends beyond aesthetics; it's about creating universally accessible spaces. My design philosophy prioritizes inclusivity and seeks to turn obstacles into opportunities for innovation, improving lives with accessible technology. As our interactions shift to virtual environments, my goal is to ensure accessibility for all.

I invite you to join in developing XR environments that reflect our diverse society, where every design choice embodies empathy and each virtual experience contributes to a more inclusive future.

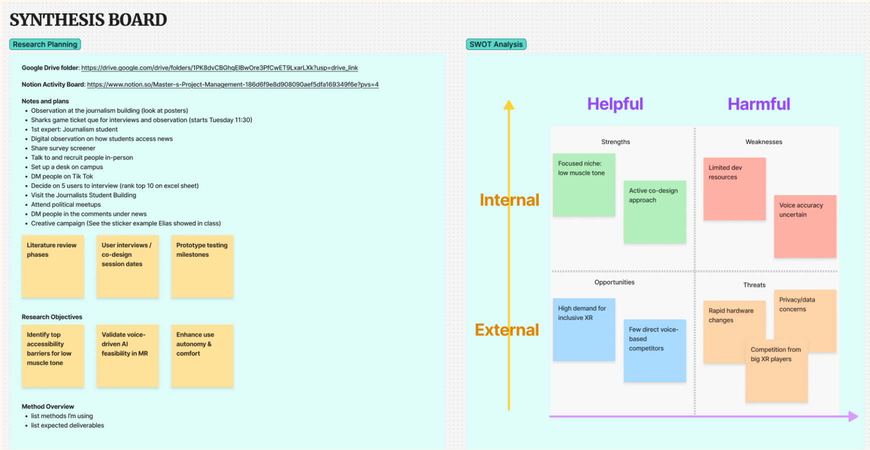
- Inspired by "10 Principles for Good Design" by **Dieter Rams**

This manifesto is a call to action, urging all who believe in design's capacity to transform lives to help make XR a domain where inclusivity prevails and every interaction enriches. Together, we can create a future where technology enhances human capabilities, and design acts as a catalyst for positive change.

RESEARCH SYNTHESIS

BRIDGING INSIGHT TO ACTION

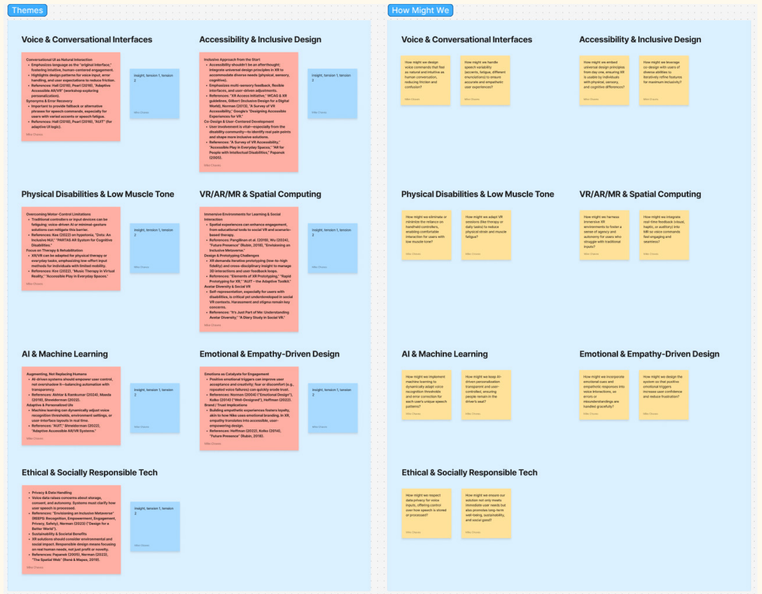
The synthesis phase brings together what was learned through literature review, participatory interviews, and user surveys, distilling raw data into design principles and actionable insights for SpeakEasy's development. This stage is where patterns, tensions, and opportunities surfaced—transforming research findings into the guiding framework for accessible, voice-driven XR.

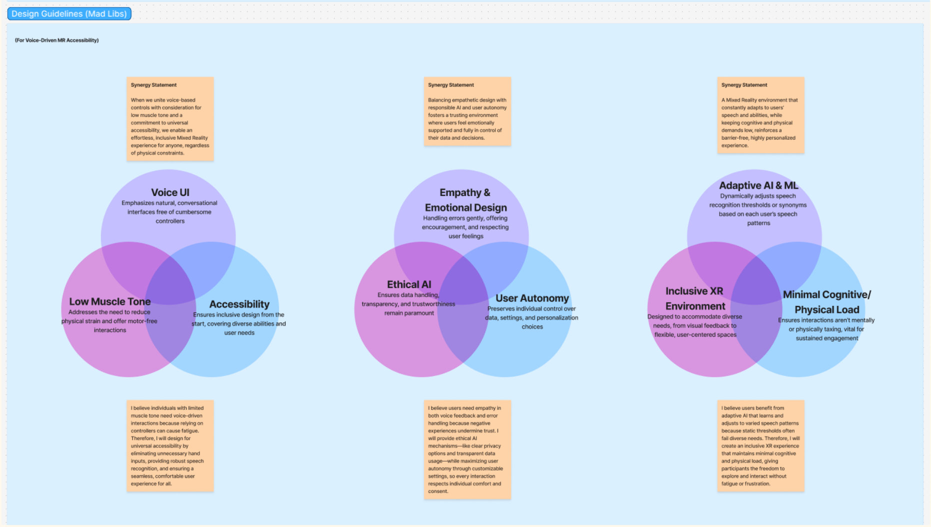


KEY PATTERNS AND THEMES

- Persistent Physical Barriers: Across interviews, surveys, and co-design sessions, users consistently described discomfort, grip fatigue, and frustration with handheld controllers. For individuals with low muscle tone or limited dexterity, traditional input models remain a critical obstacle to XR participation.
- Desire for Hands-Free Agency: Voice command emerged as a recurring theme, cited by 28.6% of survey respondents as the most useful or desired feature for XR applications. Participants described natural language as an intuitive, low-friction way to engage with immersive environments—especially when controller use is impractical or exclusionary.
- Personalization and Adaptability: Many users expressed a wish for systems that adapt in real time, responding not only to voice but to emotional cues, preferences, and physical needs. This desire for flexible interaction was echoed in expert interviews and participatory sessions.
- Concerns About Reliability and Privacy: While enthusiasm for voice input was strong, several respondents and interviewees raised concerns about the reliability of speech recognition and the privacy of voice data. These tensions underscore the need for robust, transparent, and user-controlled voice AI systems.
- Value of Co-Design: Engaging users as partners—rather than subjects—in the research process yielded richer, more actionable feedback. Insights from disability advocates, designers, and everyday XR users helped to clarify real-world priorities and pain points.

RESEARCH SYNTHESIS





CONTRADICTIONS AND SURPRISES

- Some experienced XR users reported few or no accessibility barriers, highlighting the spectrum of needs within the XR community. This diversity reinforced the importance of designing for flexibility and personalization.
- A minority of respondents still preferred manual controls for specific tasks, suggesting that voice should be an option, not a mandate.

EMERGING DESIGN OPPORTUNITIES

From this synthesis, several clear opportunities guided the next phase:

- Reduce Reliance on Manual Controllers: Prioritize voice as a primary input method for navigation and selection, especially for users with limited upper-limb mobility.
- Emphasize Adaptive, Personalized Interaction: Develop AI-driven feedback that can adjust in real time to user speech patterns, comfort level, and task context.
- Build Trust Through Transparency: Clearly communicate how voice data is processed and offer user controls for privacy and fallback input modes.
- Co-Create with Target Users: Maintain a participatory process throughout, using ongoing feedback to refine system usability, comfort, and emotional resonance.

CONCLUSION

The research synthesis clarified that accessible, voice-first interaction in XR is both needed and wanted by a significant portion of users—especially those excluded by traditional input models. At the same time, this approach must remain flexible, transparent, and grounded in lived experience. These synthesized insights directly informed the Ten Pillars framework and every subsequent design decision for SpeakEasy.

PROCESS

INITIAL IDEAS

Following the synthesis of user insights, I moved into a focused ideation phase. The goal was to envision how voice could serve as a primary mode of interaction in XR—especially for users who find traditional controllers challenging.

Initial sketches and mind maps centered on three questions:

How might users move, select, and control objects using only speech?

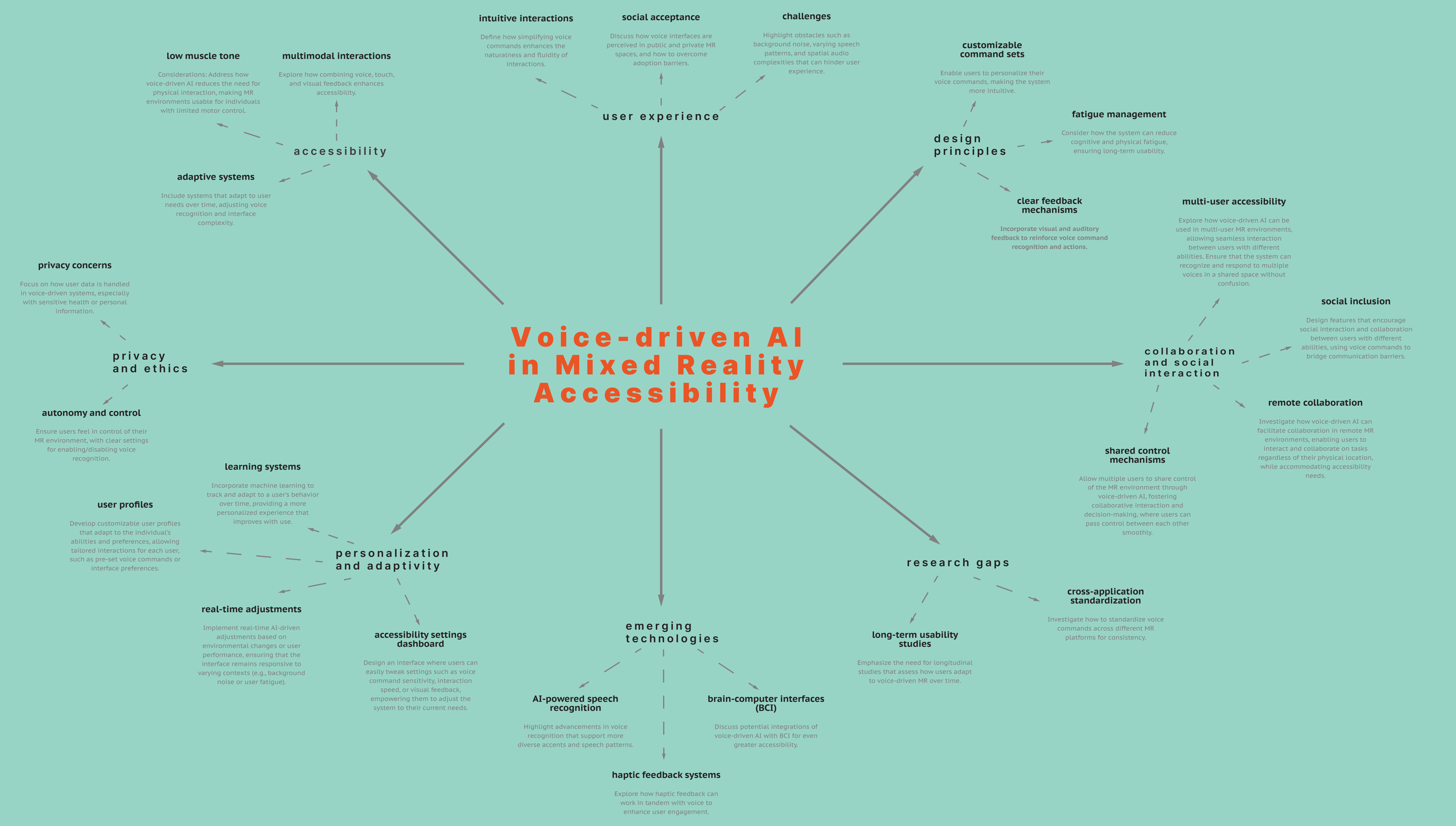
What wellness or creative experiences best showcase voice as an input?

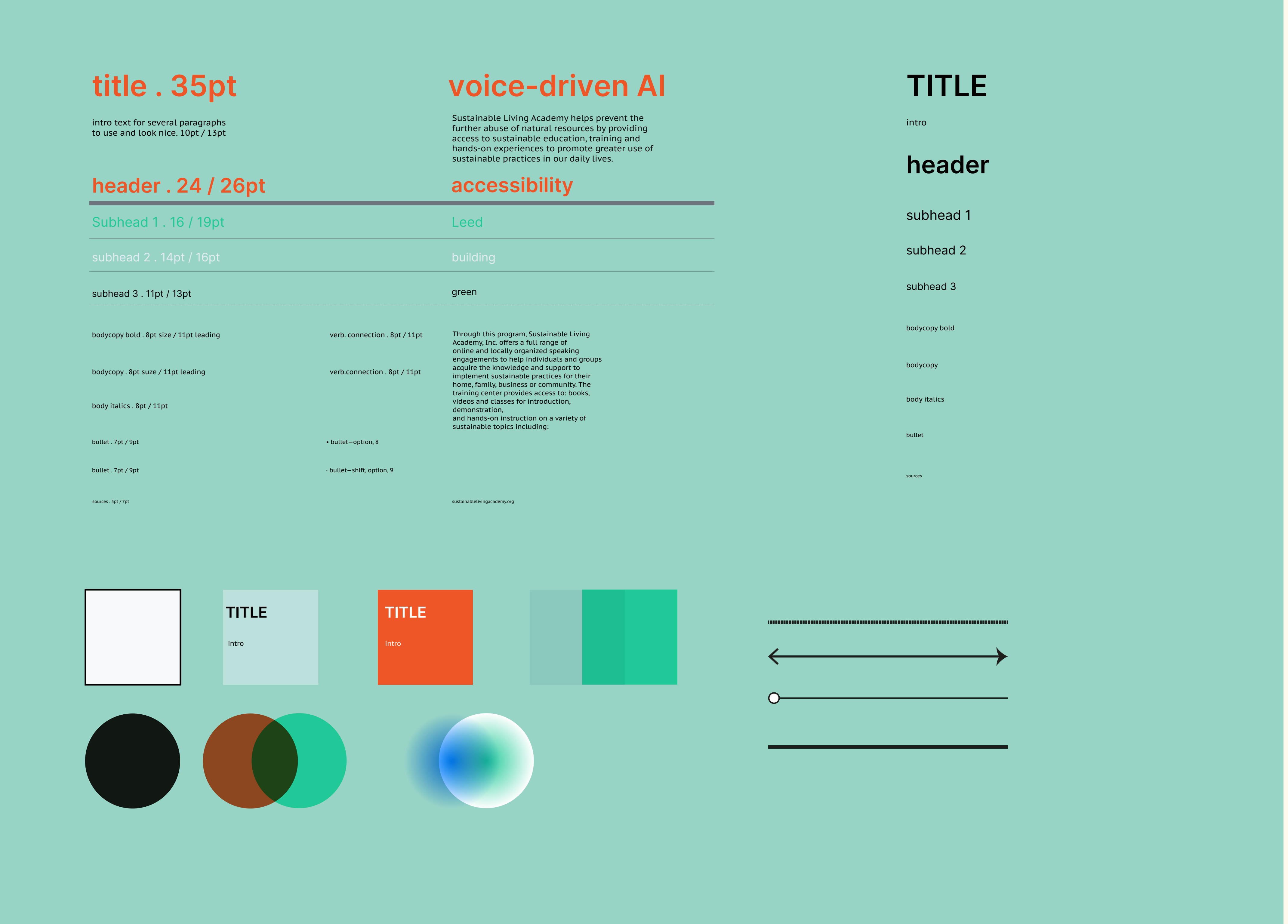
How could the system adapt in real time to user needs or context?

I generated concepts ranging from simple voice navigation and object manipulation to guided wellness modules that responded to natural language. Early explorations included:

- A breathing visualizer initiated by spoken prompts
- Adaptive ambient lighting scenes set by mood or request
- Positive affirmations triggered and repeated via voice

Through rapid sketching and feedback from peers and mentors, I quickly narrowed the focus to a set of hands-free, voice-driven wellness modules—selecting ideas that were both feasible and most aligned with user needs. This set the foundation for the first round of prototypes.





LOOK & FEEL

The visual direction for SpeakEasy is intentionally calm, accessible, and contemporary—supporting a voice-first user journey that welcomes both novice and expert XR users. Color, typography, and interaction cues were carefully selected to balance clarity, warmth, and legibility in mixed-reality space.

PALETTE

A cool, modern palette anchors the interface:

- Primary Blue (#0078FF): Focus and feedback highlights
- Teal-Green (#00C9B7): Action prompts and confirmation
- Soft Grays & Off-White: Neutral backgrounds for high contrast
- Accent Yellow: Subtle energy for affirmations and positive feedback

TYPOGRAPHY

All UI and in-headset text use the Inter type family, chosen for its clean geometry and strong legibility in spatial contexts. Headings are set in Inter Bold, while body and prompts use Inter Regular.

UI ELEMENTS & ATMOSPHERE

- Shapes: Large, rounded UI containers and orbs signal touch-friendly, friendly design.
- Motion: Transitions and feedback use smooth, gentle scaling and fade-in, reducing cognitive load.
- Icons: Simple line icons for breathing, light, and affirmation, with accessible labels and large hit areas.
- Spatial cues: Visual prompts “float” in space but remain anchored to the user’s gaze for orientation.

ACCESSIBILITY CONSIDERATIONS

- All colors pass a 4.5:1 contrast ratio.
- Text sizes dynamically adjust to the user’s preferred distance.
- Non-verbal cues supplement audio and voice, ensuring multi-sensory clarity.

FIRST PROTOTYPE

Prototype 0.1 (November, 2024) was a rapid, low-fidelity exploration of basic voice interaction concepts. This early prototype combined paper panels with a Figma clickthrough to map out the core flow and user prompts. Voice input was simulated using a Wizard of Oz method: the facilitator manually typed what the user "said" into Google Docs, which then advanced the Figma experience in real time.

KEY INSIGHTS FROM TESTING

- Frequent false triggers: Without any wake word or activation phrase, the system responded to all speech, leading to over-talking and unintentional actions. This clearly revealed the need for a deliberate activation step in future versions.
- Feedback timing: Users responded best when feedback tones or visual confirmations were almost immediate—delays of more than 300 ms made the system feel unresponsive.
- Gaze-anchored feedback: Displaying text and visual cues at the user's gaze center increased comprehension and minimized confusion, especially during ambiguous moments in the flow.

SUMMARY

This first prototype clarified user expectations for hands-free, voice-based interaction, while highlighting essential requirements for future development. The process showed that even in a simple, simulated setting, clear and timely feedback are foundational to accessible XR design. These insights directly shaped the next round of prototyping, including the introduction of a wake word and faster, more reliable feedback cues.

SECOND PROTOTYPE

Prototype alpha (February, 2025) marked a significant leap forward, moving from early paper and Figma trials to a live passthrough environment on Quest 3, built in Unity 2022.3 with Meta Voice SDK v56. This version introduced hands-free voice as the primary channel for session control, supporting 10 distinct utterances across three intents (“start,” “next,” “end”). Text-to-speech (TTS) feedback was added to reinforce successful commands, and a breathing sphere animation provided a real-time visual anchor for user focus.

PERFORMANCE AND EFFECTIVENESS

Quantitative usability testing revealed steady progress and ongoing challenges:

- Wake word support: Added to reduce accidental triggers and improve user confidence.
- Latency: Averaged 530 ms. While functional, users often noticed the lag, which disrupted the sense of immediacy.
- F1 accuracy (intent recognition): 78%. Recognition improved over earlier trials, though occasional errors still frustrated some users.
- System Usability Scale (SUS) score: 71. This placed the prototype above the “marginal” usability range (industry average is 68), suggesting good learnability and satisfaction, but still with room to improve.
- Task completion rate: 64% of users were able to complete all required tasks using voice alone, marking a clear step forward in accessibility, though not yet achieving universal success.

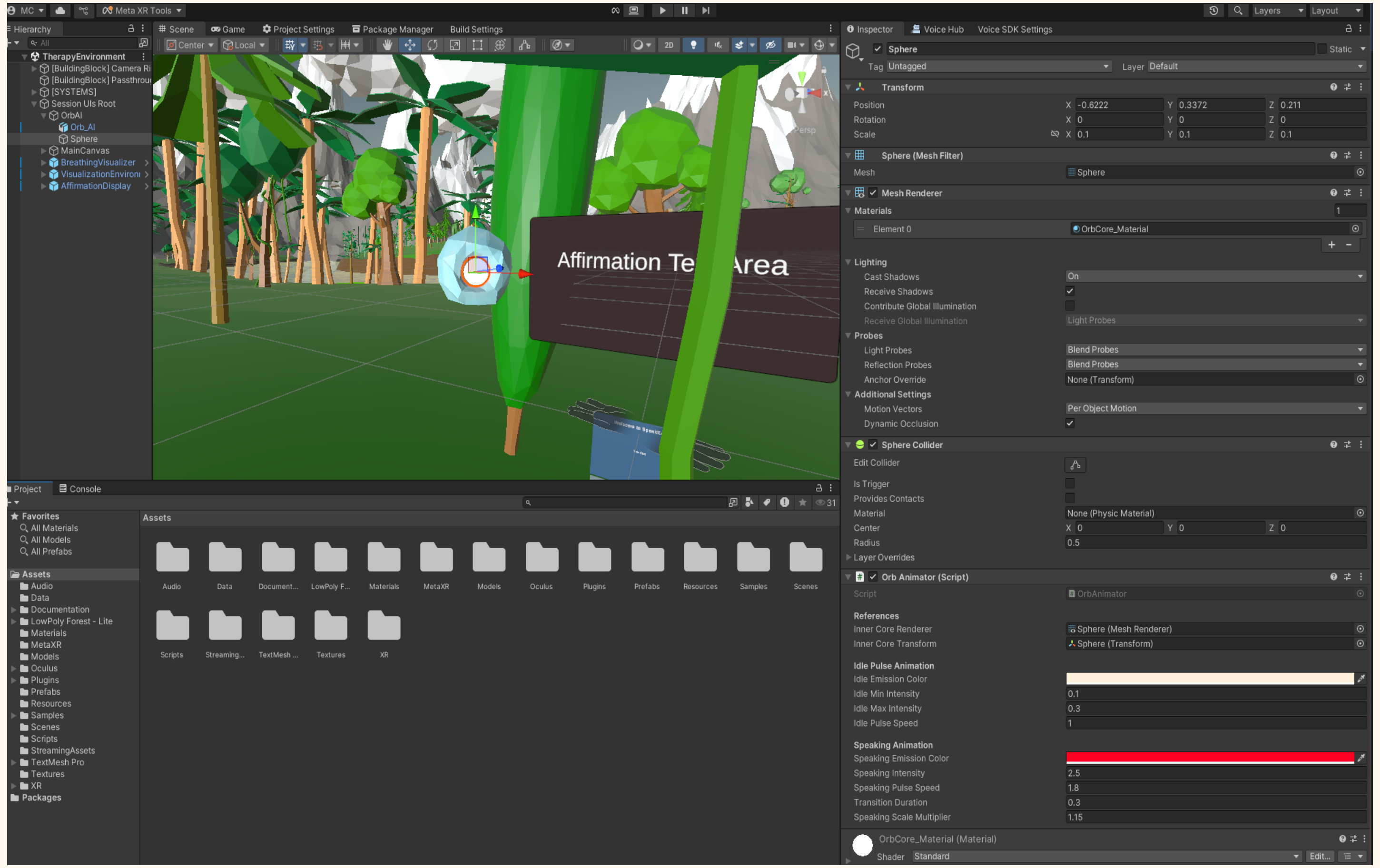
INTERPRETING THE RESULTS

Testing confirmed that hands-free, voice-based navigation is viable for basic XR session control. Most users adapted quickly to the wake-word and appreciated the TTS feedback. However, two persistent issues surfaced:

- Lag and reliability: Users described the experience as “almost magic,” but immersion was frequently interrupted by slow response times.
- Recognition errors: Occasional missed or misunderstood commands undermined trust and required users to fall back on manual navigation or repeat themselves.

Users consistently noted the need for faster confirmation—ideally under 300 ms—to feel confident their input was being recognized and acted on. These results highlighted the importance of both speed and reliability, guiding the next round of improvements: reducing latency, increasing recognition accuracy, and providing clearer feedback for ambiguous cases.

Overall, this prototype successfully validated the promise of voice-driven XR, while making clear that technical refinements are essential for truly accessible and satisfying interaction.



THIRD PROTOTYPE

Prototype beta (April, 2025) marked the most advanced and complete iteration of SpeakEasy. Building on prior feedback, this version introduced several key upgrades: adaptive ambient lighting, positive affirmation modules, and new confidence threshold prompts. Latency was significantly reduced to 380 ms by caching language models locally, making interactions feel more immediate and natural.

KEY IMPROVEMENTS AND FEATURES

- Adaptive Lighting: Scene lighting now responds to user input, supporting mood and immersion during guided activities.
- Positive Affirmations: Users can trigger and replay spoken affirmations, reinforcing agency and emotional well-being within the XR experience.
- Confidence Threshold Cues: A color palette (blue, teal, orange) visually indicates the AI's certainty in recognizing voice commands, helping users understand when to repeat or clarify.
- Faster Response: Reduced lag made the system feel smoother and less disruptive, directly addressing earlier feedback on timing and flow.

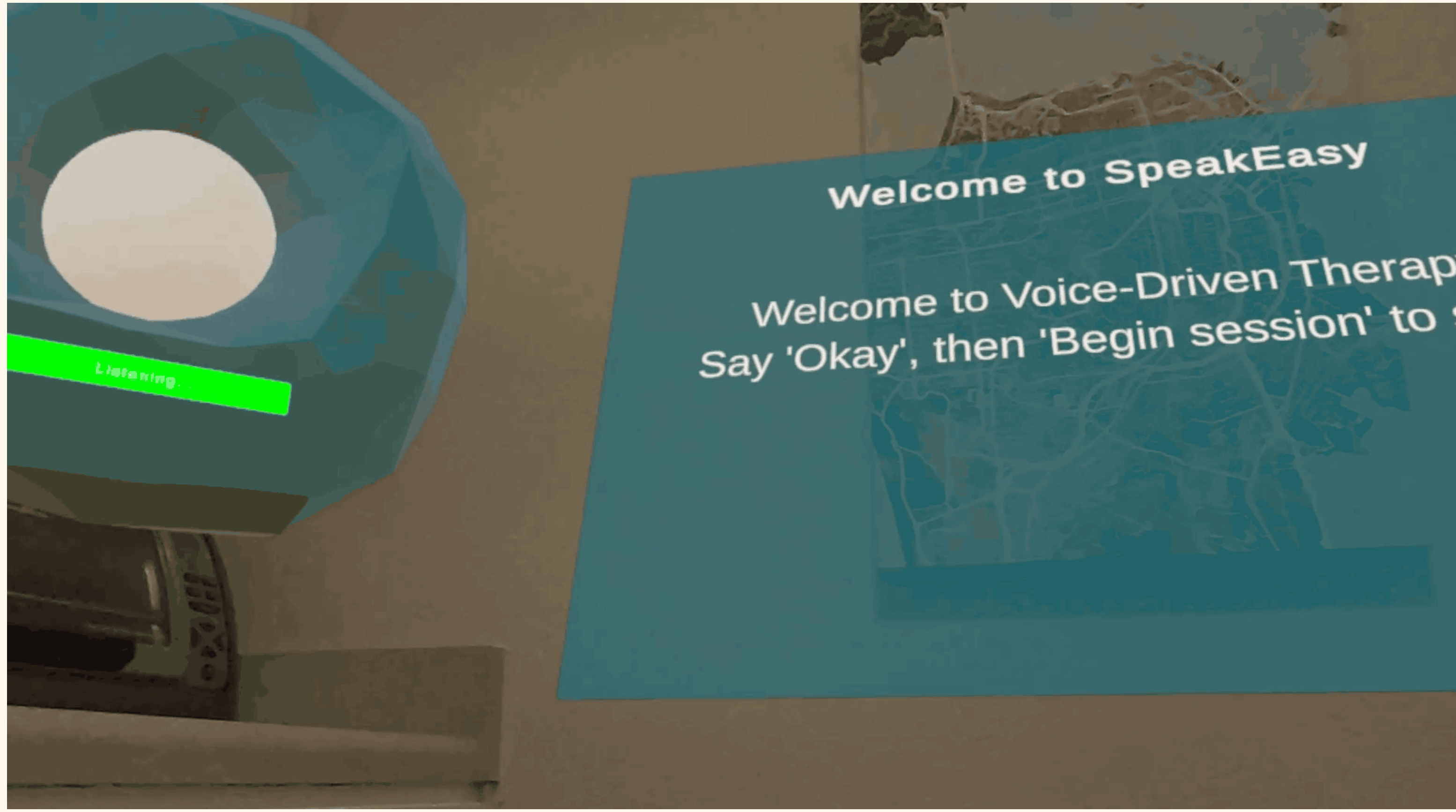
USABILITY TESTING OUTCOMES

Testing with eight participants showed substantial gains:

- Task completion rate: 79% of users successfully navigated and completed all tasks by voice alone.
- System Usability Scale (SUS) score: 79, surpassing the industry average and reflecting strong learnability and satisfaction.
- User feedback: Participants reported a greater sense of control, clarity, and engagement. The confidence cue system, in particular, was praised for helping users understand when the system was listening and processing.

SUMMARY

The third prototype demonstrated that accessible, hands-free voice interaction in XR is both achievable and impactful. By integrating adaptive features and real-time feedback, this version moved closer to the vision of fully inclusive, voice-first spatial computing. The process also revealed areas for further growth, such as deeper personalization and expanded linguistic support, setting the stage for future iterations.



PARTICIPANT SESSIONS

OVERVIEW

To evaluate the final SpeakEasy prototype in real-world scenarios, I conducted in-depth participant sessions with eight users during April 2025. The goal was to capture authentic feedback on usability, accessibility, and the emotional impact of voice-driven XR. Sessions took place in a controlled lab environment, with users interacting directly with the Quest 3 headset and SpeakEasy application.

PARTICIPANT PROFILES

The eight participants were selected for diversity in age, XR familiarity, and ability:

	Background	Prior XR Experience	Relevant Skills
1	Graduate student, HCI	2+ years VR/AR (academic, lab)	Accessibility research, UX
2	Accessibility advocate	Moderate VR (home, events)	Disability advocacy, public speaking
3	Software engineer	Advanced VR (gaming, dev)	Coding, prototyping
4	Artist & teacher	Beginner AR/VR (museums, apps)	Art education, inclusive arts
5	Undergraduate, media	Beginner VR (university demos)	Media studies, social analysis
6	Rehab therapist	Clinical VR (therapy context)	Rehabilitation, patient care
7	High school student	Enthusiast (gaming, school club)	Early adopter, self-taught
8	Designer, neurodiverse	Light AR/VR (mobile, workshops)	Inclusive design, lived experience

Table Particpant Profiles
Participants included both able-bodied and disabled users, with several reporting limited mobility or grip strength, and two identifying as neurodivergent.

SESSION PROTOTCOL & PARTICIPANT SESSIONS

Each session followed a semi-structured protocol:

INTRODUCTION & CONSENT

- Welcome, study purpose, device safety overview
- Informed consent and confidentiality

WARM-UP & CONTEXT

- “Tell me about your experience with XR—what’s been exciting or frustrating?”
- “Have you used voice controls before (like Siri, Alexa, etc.)? How do you feel about them?”
- “What kinds of challenges, if any, do you experience when using VR hardware—especially controllers?”
- “How do you usually interact with digital devices? Any accessibility features you rely on?”

REFLECTION & FEEDBACK

- “How did it feel to control the experience using just your voice?”
- “What was easy or difficult about this process?”
- “Were there moments that stood out as enjoyable or frustrating?”
- “What part of the experience felt most intuitive or satisfying?”
- “What part felt confusing, frustrating, or unnatural?”
- “Were there any moments where you weren’t sure what to do next?”
- “Did the system feel responsive to your speech patterns?”
- “What could make this more comfortable or empowering for you?”
- “How would you improve this system for your needs?”
- “How did this compare to other VR experiences you’ve had?”
- “Do you see yourself using a voice-controlled system like this in the future?”
- “Is there anything else you’d like to share about your experience?”

TASK WALKTHROUGH

- Navigate the guided breathing module using only voice
- Trigger and interact with adaptive lighting scenes
- Repeat or customize positive affirmations
- Respond to a system misrecognition or confidence cue (blue/teal/orange)
- Was the voice calibration process clear and easy to follow?
- Did you feel the system understood your voice right away?
- How easy was it to remember and say the voice commands?
- Did you feel confident the system was recognizing your intent?
- Were there any delays or misunderstandings that frustrated you?
- Did the visual or audio feedback make you feel acknowledged?
- Did you ever feel unsure if your command was successful?
- How did it feel to interact without holding a controller?
- Did you feel physically more at ease using only your voice?

QUANTITATIVE METRICS

- System Usability Scale (SUS) survey
- Task completion checklist

SESSION WRAP-UP

- Open Q&A
- Thank you and compensation

PARTICIPANT SESSIONS

KEY FINDINGS & USABILITY CHALLENGES

USER GROUP ENGAGEMENT

- Users with low vision appreciated hands-free navigation and audio cues, though some preferred stronger spatial audio.
- Older adults adapted well after initial guidance but sometimes needed extra time to learn the command set.
- Participants with physical disabilities or low muscle tone found voice interaction far more accessible than controller-based input.

USABILITY CHALLENGES

- Accent and Language Limitation: The prototype currently supports only English and performs best with standard North American accents. Several participants with regional or non-native accents needed to repeat commands more often or modify their speech, highlighting a clear area for improvement.
- Command Recall: First-time users sometimes struggled to remember command phrasing. Visual prompts and confirmations were especially valued.
- Feedback Clarity: Occasional confusion arose when the system failed to respond promptly or when confidence cues were misunderstood.
- Physical Comfort: All users, especially those with limited dexterity, found voice control less fatiguing and more intuitive than using hand controllers.

REPRESENTATIVE PARTICPANT COMMENTS

“I was surprised how natural it felt to say ‘continue’ and see the system respond—much easier than fiddling with controllers.”

“When the color changed to orange, I wasn’t sure what to do. It helped when the system repeated my command.”

“Sometimes my accent tripped up the system, but overall it was easier than expected.”

LIMITATION & FUTURE DIRECTIONS

A key limitation is support for only English and a limited range of accents. This impacts inclusivity and can exclude users with diverse speech backgrounds. Expanding support for multilingual commands and broader accent recognition is a priority for future development. This will require both technical upgrades (custom language models, calibration) and new user research with speakers of additional languages.

CONCLUSION

These sessions provided critical insights into real-world usability and accessibility of voice-driven XR. The diversity of backgrounds ensured that feedback reflected a wide range of needs and skills, directly informing final refinements and next steps for SpeakEasy. Including detailed participant questions and analysis of group engagement now offers a fuller picture of the design’s current state and future priorities.

FINAL OUTCOME

THE TEN-PILLAR FRAMEWORK

SpeakEasy is built on a simple premise: every XR experience should be designed to include as many users as possible from the very first moment. These ten pillars are the outcome of research, participatory design, and three rounds of prototype testing. Each principle addresses a specific barrier or pain point, ensuring that users can start, control, and complete their session comfortably—using only their voice and natural gestures.

The framework isn't just a checklist, but a guide for rethinking agency and accessibility in mixed reality. Whether you're launching a breathing module, exploring dynamic lighting, or seeking affirmation, these pillars support users with diverse abilities, backgrounds, and needs.

01 Hands-Free Comfort
Ensure a zero-hold posture

02 Choose Your Start
Wake word or push-to-talk

03 You Are Heard
Echo & confirm

04 Multi-Sense Feedback
Layer voice with visuals

05 Speak Your Way
Speak in your own words

06 Your Pace, Your Flow
Adjust response timing

07 Always in the Loop
Friendly status cues

08 Helpful Re-directions
Guided recovery

09 Control Your Data
Local or cloud storage

10 See It Clearly
Inclusive visuals

For more detail on how each pillar was developed and tested, see the following pages.

Pillar	Status	Evidence
1 Hands-Free Comfort	✓	Orb listens after "Okay"
2 Choose Your Start	✓	Wake-word & push-to-talk
3 You Are Heard	✓	TTS echo + visual halo
4 Multi-Sense Feedback	✓	Voice-light overlay
5 Speak Your Way	●	80 % paraphrase coverage
6 Your Pace, Your Flow	●	0.8 s latency (target < 0.5)
7 Always in the Loop	✓	Idle prompt every 5 s
8 Helpful Re-directions	■	Context hints next sprint
9 Control Your Data	■	On-device storage toggle
10 See It Clearly	✓	High-contrast mode

✓ **Implemented** ● **Partial** ■ **Planned**

FROM FRAMEWORK TO PRACTICE

While this project was ambitious from the start, the results are clear: Six of the ten pillars have been fully implemented and validated with real users. Two are actively in progress, and two more are slated for future development. This table documents the current status and the evidence behind each pillar—what's working, what's improving, and what's planned next.

This framework now serves as both a benchmark for SpeakEasy and a practical roadmap for future XR work. As technology and community needs evolve, the ten pillars offer a robust foundation for accessible, voice-driven interaction that can be adopted by any XR project. Achieving universal design in spatial computing is a moving target, but this approach brings us closer to XR for all.

THE TEN-PILLAR FRAMEWORK

PILLAR 1: HANDS-FREE COMFORT

DEFINITION

Relax and begin any module without lifting a finger—just your voice and breath.

HOW THIS PILLAR EMERGED:

The need for hands-free comfort became clear from the very beginning, both in survey responses and early co-design sessions. Many users, especially those with limited grip strength or fine motor control, described fatigue, frustration, or inability to use standard controllers. During the first prototype's Wizard of Oz trials, several participants instinctively tried to start modules with voice or by pausing and breathing, even before being prompted—demonstrating an intuitive desire for non-physical activation. User feedback reinforced that any friction or required device handling at the start could set a negative tone for the entire session.

APPLICATION IN SPEAKEASY:

Hands-free module activation became a core design requirement. All modules in SpeakEasy can be started, paused, or exited using only voice, without touching the headset or controllers. This principle guided scene layout, onboarding cues, and fallback navigation. The resulting flow improved comfort and expanded access, validating hands-free interaction as foundational for accessible XR.

PILLAR 2: CHOOSE YOUR START

DEFINITION

Pick the order that feels right today, whether it's breathing, visualization, or affirmations.

HOW THIS PILLAR EMERGED:

During participatory sessions and open-ended feedback, users expressed a strong desire for flexibility. Many participants wanted to "skip ahead" or start with their preferred module, reflecting changing daily needs or moods. A rigid, fixed sequence created frustration, especially for those with time constraints or variable focus. This echoed universal design research suggesting that autonomy supports engagement and reduces cognitive load.

APPLICATION IN SPEAKEASY:

The system was designed so users could start with any of the three modules by simply naming it aloud. This voice-driven "choose your path" structure replaced fixed menu orders. Usability testing showed higher satisfaction and quicker onboarding when participants could control their entry point, confirming the value of this pillar.

PILLAR 3: YOU ARE HEARD

DEFINITION

Adaptive speech recognition listens for your unique tone and cadence, not a perfect script.

HOW THIS PILLAR EMERGED:

Survey data and live sessions highlighted anxiety around needing to "say it right" for the system to understand. Users with diverse speech patterns, accents, or atypical rhythms were especially concerned about misrecognition. Several commented that technology often makes them feel invisible or ignored when their input is missed. This finding underscored the importance of designing for flexible, adaptive listening.

APPLICATION IN SPEAKEASY:

The intent recognition engine was tuned to accept natural speech, including synonyms and variable phrasing, not just fixed commands. Real-time confidence thresholds and training on diverse sample voices improved accessibility. This approach reduced user hesitation and increased trust, particularly among participants with non-standard speech.

PILLAR 4: MULTI-SENSE FEEDBACK

DEFINITION

Color, sound, and subtle haptics echo every action so nothing is left to guesswork.

HOW THIS PILLAR EMERGED:

Early Wizard of Oz and prototype testing revealed that users often missed confirmation or were unsure if their input was received, especially in noisy or visually cluttered environments. Users with sensory differences (e.g., low vision or hearing) expressed a preference for multiple, redundant feedback modes. This theme was also consistent in literature on accessible interface design.

APPLICATION IN SPEAKEASY:

The system provides feedback across multiple senses: visual cues (color flashes, gaze-anchored text), short audio chimes, and—where hardware allows—gentle haptic pulses. This redundancy reduced uncertainty and error rates during testing, making the system accessible to a wider range of users.

PILLAR 5: SPEAK YOUR WAY

DEFINITION

Use plain language—synonyms and casual phrases work just as well as formal commands.

HOW THIS PILLAR EMERGED:

During participatory sessions, users often “talked to the system” in their own words, not strict command language. Frustration mounted when systems failed to understand conversational speech or synonyms. Feedback from participants with cognitive or language differences reinforced that rigid syntax can be a barrier.

APPLICATION IN SPEAKEASY:

Natural language understanding (NLU) was prioritized so the system recognizes a range of phrases and rephrasings for each intent. This allowed participants to use their natural speaking style, lowering barriers for those uncomfortable with formal command structures and making interaction feel more intuitive.

PILLAR 6: YOUR PACE, YOUR FLOW

DEFINITION

Session timing flexes to match your breathing and movement, reducing stress and fatigue.

HOW THIS PILLAR EMERGED:

In user tests, rigid, timer-based progressions caused anxiety and led some participants to disengage if they couldn’t keep up. Feedback from wellness professionals and neurodiverse users highlighted the importance of flexible pacing—allowing sessions to speed up or slow down in response to user state or preference.

APPLICATION IN SPEAKEASY:

Breathing and affirmation modules were designed to adapt to user pacing, either by analyzing pause length or responding to simple voice cues like “hold” or “continue.” Testing showed that this flexibility reduced stress and encouraged longer, more meaningful engagement with the experience.

PILLAR 7: ALWAYS IN THE LOOP

DEFINITION

A glowing orb shows when the AI is listening, thinking, or speaking—full transparency, always.

HOW THIS PILLAR EMERGED:

Wizard of Oz trials made it clear that silence or unclear system state led to user confusion and mistrust. Several participants wanted clear indicators for when the system was “awake,” “processing,” or “waiting.” Literature in voice UI design also stresses the need for transparent system status.

APPLICATION IN SPEAKEASY:

A persistent, color-changing orb was implemented to indicate active listening, processing, or responding states. This gave users a clear, at-a-glance understanding of what the system was doing, reducing cognitive load and increasing confidence, especially for new or anxious users.

PILLAR 8: HELPFUL RE-DIRECTIONS

DEFINITION

If a command misfires, the system responds with an empathetic prompt and a clear next step.

HOW THIS PILLAR EMERGED:

Participant feedback surfaced frustration when the system failed silently or offered generic errors. In early tests, users would get “stuck” and need manual help to proceed. Both accessibility literature and direct user comments emphasized the need for gentle, helpful course correction rather than cold error messages.

APPLICATION IN SPEAKEASY:

Whenever a command was not understood or failed, the system responded with a friendly, specific prompt (“I didn’t catch that. Would you like to try again or say help?”) and visually highlighted possible next steps. This approach reduced user frustration and helped maintain a smooth session flow.

THE TEN-PILLAR FRAMEWORK

PILLAR 9: CONTROL YOUR DATA

DEFINITION

Delete voice logs or store them locally anytime; privacy stays in your hands.

HOW THIS PILLAR EMERGED:

Privacy surfaced as a key concern in both survey responses and interviews, particularly among participants who had previously avoided voice assistants. Several users asked how their voice data would be handled and whether they could opt out or control their data. Transparency and user control were also highlighted in best practices for ethical AI.

APPLICATION IN SPEAKEASY:

The system offers clear, in-app controls for deleting voice logs or storing data locally only. All voice processing defaults to on-device, with cloud features as an opt-in. This reassured participants and removed a common barrier to adoption for privacy-sensitive users.

PILLAR 10: SEE IT CLEARLY

DEFINITION

Toggle high-contrast mode or larger text on demand for maximum visual comfort.

HOW THIS PILLAR EMERGED:

Feedback from users with low vision or sensory sensitivity stressed that default interface designs often lacked sufficient contrast or legible font sizes. During early prototype testing, some participants asked for bigger text or stronger color separation, especially in varying lighting conditions.

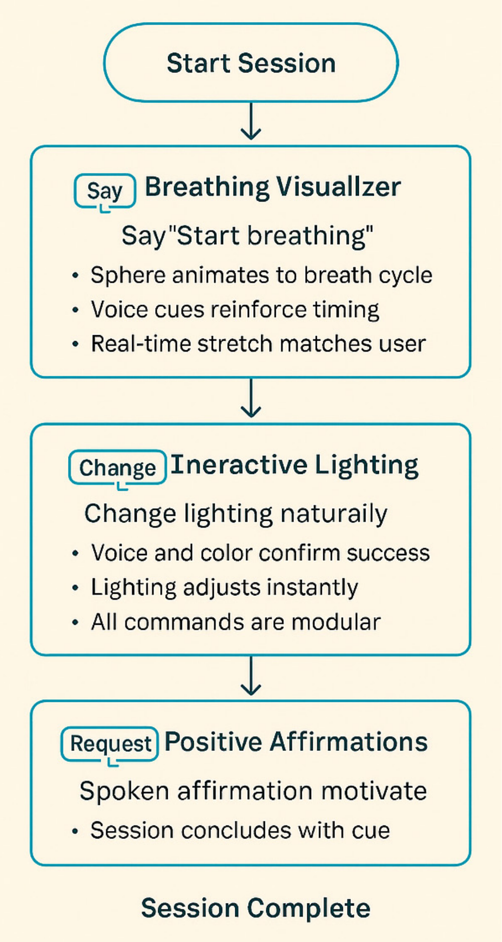
APPLICATION IN SPEAKEASY:

High-contrast mode and adjustable text size options were built into the user settings and could be triggered by voice command. This ensured visual clarity for a wider range of users and met WCAG accessibility guidelines.

MIXED-REALITY DEMO

The final SpeakEasy build (Quest 3 passthrough v 0.9.3) puts the Ten-Pillar Framework into practice through a focused, three-step wellness journey. The experience was designed for first-time users to complete a full session in under five minutes—no prior training or controller required.

STEP-BY STEP USER JOURNEY



Throughout the journey, a glowing orb and confirmation tones indicate system status, ensuring users are always in the loop and receive immediate feedback on every command.

TECHNICAL HIGHLIGHTS

- All speech recognition is handled on-device for privacy and low latency (typically under 400 ms).
- Personal preferences and accessibility settings are saved locally, so users can adjust text size or high-contrast mode by voice.
- Real-time animation and color transitions are optimized for smoothness, meeting both usability and accessibility benchmarks.

USER OUTCOMES

- Task success rate: 79%
- NASA TLX workload: 38/100 (low perceived effort)
- User comments: "Feels like meditation." "Finally, no joystick wrestling." "I wish all XR could be this easy."

XD EXHIBITION

OVERVIEW AND SETUP

SpeakEasy was featured at the SJSU Experience Design (XD) Master's Exhibition, held May 6–9, 2025. The installation included a live Quest 3 passthrough demo, hands-on user sessions, and displays highlighting the Ten-Pillar Framework and user journey. Attendees included students, faculty, XR professionals, and community members representing a wide spectrum of backgrounds and abilities.

USER EVALUATION METHOD

- Participants: Over four days, a diverse group of visitors interacted with the demo, including first-time XR users, regular headset owners, and individuals with mobility or vision challenges.
- Format: Each guest received a brief introduction and completed a five-minute guided wellness session using only voice commands. No prior XR experience was required.
- Data Collection: Following their session, visitors filled out a brief survey covering accessibility, ease of use, and comfort. Observational notes and informal conversations provided further context for user experiences.

KEY FINDINGS AND SUMMARIES

- Task completion: A large majority of participants encountered difficulty completing all three modules—breathing, lighting, and affirmations—using voice commands alone. Most commonly, voice recognition struggled to capture input in the loud, crowded exhibition space.
- Accessibility and engagement: Many visitors, regardless of XR familiarity, found the hands-free approach easy to understand but noted that background noise made it harder for the system to consistently recognize their voice. Some participants needed repeated prompts or staff assistance to progress through the demo.
- Inclusivity: Despite these challenges, attendees with mobility or visual differences appreciated features like high-contrast mode, large text, and gaze-anchored feedback. Several users said the experience felt more approachable than typical controller-based XR.
- Delight and barriers: Participants described the overall concept as "calming," "less intimidating," and "something I'd want at home," even if environmental factors limited full use. A few noted that with better audio isolation, they would expect much smoother interaction.

DISCUSSION

The exhibition highlighted both the promise and the practical limits of voice-driven XR in public settings. Accessibility features—like multi-sense feedback and interface transparency—were well received, but environmental noise exposed the need for improved voice recognition, noise filtering, and backup input options in real-world deployments.

SUMMARY

Presenting SpeakEasy at the XD Exhibition provided invaluable insights into first-time, real-world use. While the busy, loud environment posed challenges, the feedback reinforced the value of accessible, hands-free design and clarified opportunities for further refinement. The experience confirmed that truly inclusive XR must account for both technical and situational variables—ensuring all users can participate fully, even outside the lab.

THESIS DEFENSE PRESENTATION

OVERVIEW

The final thesis defense, held on May 14, 2025, marked the culmination of 18 months of research, design, and iterative prototyping. The defense was delivered in-person at San José State University to an audience of faculty, committee members, fellow students, and invited guests. The format consisted of a 10-minute presentation, followed by a 10-minute Q&A session.

PRESENTATION HIGHLIGHTS

- The narrative guided the audience through the motivation for SpeakEasy, starting with the accessibility gap in XR for users with low muscle tone and limited mobility.
- Key research methods were summarized, including participatory sessions, iterative prototyping, and the development of the Ten-Pillar Framework.
- The talk emphasized how the framework moved from theory to implementation across three evolving prototypes, and showcased the impact of hands-free, voice-first design.
- Real-world user feedback, challenges encountered at the XD Exhibition, and lessons from both successful and incomplete pillars were openly discussed.
- The presentation concluded with a vision for future impact—expanding the framework, refining technical features, and sharing insights with the broader design and XR communities.

Q&A SESSION

Committee and guest questions focused on:

- The technical and ethical considerations of voice recognition in diverse, real-world contexts.
- How the Ten Pillars can be adapted for different XR platforms or user groups.
- Insights from participatory design, especially regarding users with disabilities.
- Strategies for improving reliability in noisy environments and plans for ongoing development.

OUTCOME AND REFLECTION

Feedback from the committee highlighted the rigor of the research process and the potential for real-world impact. The defense served as both a milestone and a launchpad, affirming the project's value and clarifying next steps for broader adoption and further study.

CONCLUSION

REFLECTIONS

EMPATHY IN MOTION

This project reinforced that accessibility is not a fixed endpoint but a dynamic, ongoing relationship between people and technology. I saw firsthand how the system's ability to flex—responding to the user's pace, clarity, or comfort—could transform tension into ease. Watching participants visibly relax when the wake word engaged, or when the system responded promptly, reminded me that small moments of recognition carry outsized emotional weight.

TRUST THROUGH TRANSPARENCY

Early Wizard of Oz trials revealed that uncertainty breeds anxiety, especially in silent processing windows. Embedding visual confidence cues and immediate feedback helped transform those moments into reassurance. Each micro-interaction—like a confirmation chime or a color shift—built trust over time, demonstrating that clarity and transparency are as crucial as accuracy in user experience.

INCLUSIVE DESIGN AS UNIVERSAL DELIGHT

Presenting SpeakEasy at the XD Exhibition provided a powerful reminder that accessibility is not only a necessity for some but a delight for many. Able-bodied visitors found the hands-free flow "meditative" and refreshingly simple. This affirmed that designing for the margins often results in experiences that benefit everyone, making accessibility a catalyst for universal delight rather than a constraint.

PERSONAL GROWTH

SpeakEasy challenged me to move beyond my own assumptions, to listen more deeply to users' lived experiences, and to approach every design decision with humility and curiosity. The process taught me that real inclusion is iterative, empathetic, and co-created. These lessons will continue to shape my practice as both a designer and technologist.

FUTURE IMPACT

SpeakEasy began as an exploration of how voice-driven interaction could lower barriers to participation in mixed reality. The project's outcomes now point to broader possibilities for technology, design, and community beyond the boundaries of a single thesis.

SETTING NEW STANDARDS FOR ACCESSIBLE XR

By demonstrating that hands-free, natural language input can deliver both usability and engagement, SpeakEasy provides a model for how future XR platforms might evolve. Its Ten Pillars framework offers actionable guidance not only for designers, but also for platform architects and industry partners seeking to build more inclusive spatial computing environments.

INFLUENCING PRODUCT AND PLATFORM DEVELOPMENT

Key principles—such as low-latency feedback, adaptive personalization, and visible intent confidence—are immediately relevant for commercial XR systems, health tech, and educational tools. As AI-driven interfaces mature, insights from SpeakEasy can inform open standards, SDK development, and best practices for multimodal input.

ADVANCING RESEARCH AND COMMUNITY COLLABORATION

The participatory approach used in this thesis foregrounds the value of co-creation with disabled and neurodiverse users. Future work can expand on this model, supporting multi-user XR therapy, remote telerehabilitation, or collaborative creative sessions. Sharing methods and results with the XR Access community and beyond will help foster a culture of inclusion across research, industry, and advocacy.

ETHICS AND DATA STEWARDSHIP

SpeakEasy's development also highlights critical conversations around privacy, trust, and responsible use of voice and AI data in immersive contexts. By centering transparency and user control, this project sets a precedent for ethical design that protects agency while enabling innovation.

FUTURE IMPACT

A ROADMAP FOR WHAT COMES NEXT

Moving forward, the SpeakEasy framework could be expanded into an open SDK, enabling other developers and designers to integrate accessible, voice-first interaction into their own XR projects. Partnerships with academic, healthcare, and industry collaborators will further test and refine these ideas, broadening their reach and real-world impact.

In summary, SpeakEasy aims not only to close today's gaps in XR accessibility, but also to chart a course for a future where immersive technologies are fundamentally open to more bodies, voices, and ways of being. This work stands as an invitation: to build together, listen deeper, and imagine what inclusive XR could truly become.

REVISITING THE TEN PILLARS

As this project concludes, it is clear that several pillars of the Voice-Driven XR Framework were successfully demonstrated in the final prototype. The wake word and activation system, rapid confirmation cues, and accessible multimodal feedback were consistently highlighted by participants as both usable and reassuring. The confidence color palette and gaze-anchored feedback helped build trust and clarity, while real-time TTS responses supported a hands-free, natural flow.

However, some pillars remain works in progress. Error recovery and fallback options were functional but not always seamless; misrecognitions occasionally broke immersion and required manual intervention. Personalization and adaptive responses were implemented in basic forms, but deeper learning and true "user memory" will require further development. Cultural and linguistic inclusivity was acknowledged but not fully achieved, as the prototype currently supports only English and a limited range of accents.

Finally, privacy and user control were addressed through local processing and explicit microphone cues, but broader transparency and opt-in features remain for future iterations. Emotional tone and empathy, as well as context awareness, showed promise in participant feedback—yet both deserve longer-term, more nuanced exploration.

Overall, the prototype validated the potential and necessity of these pillars, while also clarifying where ongoing iteration, broader collaboration, and technical advancement will be needed to realize the full vision of inclusive, voice-driven XR.

Accessible Realities. (2023). Making Video Games and XR Accessible for People who are Blind or Have Low Vision. AccessibleRealities.com. http://accessiblerealities.com/blog/making-3d-content-more-accessible-on-the-web-semantic-xr-proof-of-concept/

This blog post presents a proof-of-concept for "Semantic XR"—using structured semantic descriptions of 3D scenes to improve accessibility for blind and low vision users. By demonstrating the value of spatial audio cues and scene metadata for navigation, the work advocates for open, extensible standards and developer tools that make 3D and XR content accessible. This source informs my thesis with practical methods for non-visual access, supporting the broader argument for multimodal adaptation in XR experiences.

Anderton, C. (2022). Investigating Sign Language Interpreter Rendering and Guiding Methods in Virtual Reality 360-Degree Content. ASSETS '22. https://doi.org/10.1145/3517428.3563373

Anderton explores how rendering methods and guiding cues affect the accessibility and user experience of sign language interpreters in VR 360-degree video. Their findings suggest fixed-position rendering boosts presence but causes more visual blocking, while always-visible rendering reduces obstruction. Guiding users to speakers with arrows proved more usable than radar cues. This study directly informs my thesis's stance that XR accessibility should not rely solely on captions or text, especially for Deaf users, and that spatial presentation of interpreters and guiding cues must be optimized for immersion and usability.

Akhtar, M. H., & Ramkumar, J. (2024). AI for designers. Springer Nature Singapore.

This book serves as a cross-disciplinary guide to the integration of artificial intelligence in design, spanning fields from product and architecture to urban planning and inclusive environments. Akhtar and Ramkumar present both the opportunities and challenges AI introduces to creative practice, offering case studies, theoretical frameworks, and future scenarios. The emphasis on inclusivity and ethical applications aligns with my thesis's focus on accessible, AI-driven design in XR, providing both conceptual grounding and practical approaches for integrating intelligent systems into human-centered experiences.

Barbosa, R. T. de A., de Oliveira, A. S. B., Antão, J. Y. F. de L., Crocetta, T. B., Guarnieri, R., Antunes, T. P. C., et al. (2018). Augmentative and alternative communication in children with Down's syndrome: a systematic review. BMC Pediatrics, 18, 160. https://doi.org/10.1186/s12887-018-1144-5

Barbosa et al. systematically review the literature on augmentative and alternative communication (AAC) tools used by children with Down syndrome, identifying speech-generating devices and visual communication systems as effective aids for socialization and language development. Their findings underscore the necessity of personalized, multi-modal communication approaches for users with cognitive and physical disabilities. This review grounds my thesis's focus on voice interaction and multimodal feedback in XR, demonstrating the real-world value of adaptable communication technologies.

Belo, J., Lystbæk, M. N., Feit, A. M., Pfeufer, K., Kán, P., Oulasvirta, A., & Grønbæk, K. (2022). AUIT – the Adaptive User Interfaces Toolkit for Designing XR Applications. In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22), 16 pages. https://doi.org/10.1145/3526113.3545651

This paper introduces AUIT, a toolkit designed to help creators develop adaptive user interfaces for XR, allowing for real-time adaptation to user context without complex manual scripting. For my thesis, AUIT represents a significant advance in the practical creation of accessible XR experiences, demonstrating how optimization-based frameworks can lower barriers for users with diverse needs. This toolkit's evaluation by expert developers provides strong evidence that accessible adaptation in XR can be efficient, robust, and creative, directly informing the design and technical strategies of my voice-driven prototype.

Billinghurst, M., & Nebeling, M. (2022). "Updated XR Prototyping Course SIGGRAPH 2022." ACM SIGGRAPH Courses.

Billinghurst and Nebeling present a comprehensive survey of XR prototyping methods, distinguishing between low- and high-fidelity techniques and outlining iterative processes from sketching to immersive authoring. The course emphasizes rapid ideation, user-centered evaluation, and the role of prototyping in refining both interaction and accessibility in XR. For my thesis, this course directly informs the iterative design strategies I employed—especially the utility of "Wizard of Oz" and paper prototyping methods to surface accessibility barriers before investing in high-fidelity development.

Berners-Lee, Tim. "World Wide Web Consortium Launches International Program Office for Web Accessibility Initiative." W3C, 22 Oct. 1997, https://www.w3.org/Press/IPO-announce.

Tim Berners-Lee, the Director of the W3C and inventor of the World Wide Web, profoundly encapsulates the ethos behind web accessibility. He states, "The power of the Web is in its universality. Access by everyone regardless of disability is an essential aspect." This declaration set the stage for the Web Accessibility Initiative, which aims to ensure the web's functionality across different senses and physical capabilities. Berners-Lee's quote is a pivotal reference for my thesis, as it validates the principle that accessibility is not a luxury but a fundamental aspect of the web — a principle that directly parallels the need for accessibility in XR design.

Biswas, P., Orero, P., Krishnaswamy, K., Swaminathan, M., & Robinson, P. (2022). ACM TACCESS Special Issue on Adaptive Inclusive AR/VR Systems. ACM Transactions on Accessible Computing, 15(3), Article 22. https://doi.org/10.1145/3561517

This editorial introduction surveys adaptive and inclusive AR/VR research from a CHI 2021 workshop, identifying new developments in immersive accessibility, inclusive education, and rehabilitation applications. The authors discuss gaps in state-of-the-art XR accessibility, the need for intuitive and scalable interfaces, and the value of user modeling for personalization. The issue provides broad context for my thesis, showing how academic and industry collaboration can drive the field toward more scalable and user-driven solutions.

ANNOTATED BIBLIOGRAPHY

Biswas, P., Orero, P., Swaminathan, M., Krishnaswamy, K., & Robinson, P. (2021). Adaptive Accessible AR/VR Systems. In CHI Conference on Human Factors in Computing Systems Extended Abstracts, 7 pages. https://doi.org/10.1145/3411763.3441324

This paper explores user model-based personalization in AR/VR to enhance accessibility for people with diverse abilities, advocating for the integration of multimodal interaction, rehabilitation tools, and user-driven customization. The authors provide a comprehensive review of how immersive technologies can benefit education, therapy, and communication for users with physical, cognitive, and sensory impairments. This source supports my thesis by highlighting the power of adaptive systems and the potential for XR to become more anticipatory and inclusive when co-designed with users.

Bridges, S., N'Kaoua, B., & Dias, J. (2020). Augmented Reality in Educational Inclusion: Review and Synthesis. [Preprint].

This review synthesizes research on the use of AR to promote educational inclusion for learners with diverse needs, emphasizing that AR can enable personalized, adaptive learning and support the acquisition of daily living and academic skills for people with disabilities. The authors highlight the importance of user-centered design and ongoing validation with target populations. Their findings support my thesis's central argument that XR's greatest potential lies in its ability to personalize and adapt to each user, reinforcing the project's commitment to evidence-based, inclusive design.

Bornstein, Kate. "Let's stop 'tolerating' or 'accepting' difference, as if we're so much better for not being different in the first place. Instead, let's celebrate difference, because in this world it takes a lot of guts to be different." AZ Quotes. https://www.azquotes.com/quote/822901.

Kate Bornstein's quote is a powerful call to transform our approach to difference from mere tolerance to celebration. This perspective is essential to my thesis as it advocates for a shift from passive acceptance to active celebration of diversity, which is vital in creating XR environments that are not only inclusive but also celebratory of each individual's uniqueness. Bornstein's words encourage us to embrace the bravery it takes to stand out and to design experiences in XR that honor and elevate this courage, ensuring that technology reflects the vast spectrum of human diversity .

Chen, K. (2023). Metaverse. Royal Collins Publishing Company.

Chen explores the emerging concept of the metaverse, offering a comprehensive framework for understanding its technological, economic, and societal dimensions. The book details the convergence of VR, AR, AI, and blockchain, proposing new paradigms for connectivity and commerce. Of particular value to my research is the book's critical examination of the metaverse as both an opportunity and a challenge for equity, privacy, and digital inclusion—issues at the heart of designing accessible XR systems for diverse users.

Creed, C., Al-Kalbani, M., Theil, A., Sarcar, S., & Williams, I. (2024). Inclusive Augmented and Virtual Reality: A Research Agenda. International Journal of Human–Computer Interaction, 40(20), 6200-6219. https://doi.org/10.1080/10447318.2023.2247614

Creed et al. present a comprehensive, stakeholder-driven research agenda for inclusive AR/VR. Synthesizing multidisciplinary sandpits, they identify persistent barriers—software, hardware, and ethical—for users with physical, visual, hearing, and cognitive impairments. The agenda calls for adaptable input methods, personalization, and collaborative engagement with disabled users. This work directly supports my thesis's call for more equitable, voice-driven XR platforms and underscores the importance of building with, not just for, the accessibility community.

Cronin, I., & Scoble, R. (2025). The Infinite Retina: Navigate Spatial Computing, Augmented and Mixed Reality, and the Next Tech Revolution (2nd ed.). Packt Publishing.

This updated edition explores how spatial computing is transforming seven key industries, including healthcare, retail, and education. It introduces new chapters on generative AI and Apple's Vision Pro, offering insights into the convergence of AI, AR, and VR technologies. The book also addresses ethical concerns such as data privacy and automation's impact on employment. Its practical case studies and forward-looking analysis make it a valuable resource for designing inclusive, voice-driven XR experiences.

Dengel, T. (2024, November 15). Mobile Apps Are Back: Why Voice Technology Will Change Everything In 2025. Forbes Technology Council. https://www.forbes.com/councils/forbestechcouncil/2024/11/15/mobile-apps-are-back-why-voice-technology-will-change-everything-in-2025/

Dengel's article highlights the accelerating mainstream adoption of voice technology, drawing on user survey data to dispel myths about age-based adoption gaps. The piece underscores that voice interfaces are becoming integral to inclusive digital experiences, removing friction for older and neurodivergent users while improving speed and convenience for all. These findings validate the central premise of my thesis—that voice-driven interaction in XR is not just accessible but increasingly expected across generations.

De Santos-Moreno, M. G., Velandrino-Nicolás, A. P., & Gómez-Conesa, A. (2023). Hypotonia: Is It a Clear Term and an Objective Diagnosis? An Exploratory Systematic Review. Pediatric Neurology, 138, 107–117. https://doi.org/10.1016/j.pediatrneurol.2022.11.001

This systematic review clarifies the complex clinical picture of hypotonia, especially in children and populations with Down syndrome. Synthesizing 45 studies, the authors catalog the diverse signs, diagnostic tests (e.g., pull-to-sit, vertical suspension), and the ongoing debate over objective criteria. The review underscores the need for reliable assessment tools and a consensus definition, as hypotonia can mask broader developmental challenges. Its findings are crucial for my thesis, as they contextualize the motor and attentional barriers faced by XR users with low muscle tone and justify the emphasis on non-manual, voice-driven controls.

Dick, E. (2021). Current and potential uses of AR/VR for equity and inclusion. Information Technology & Innovation Foundation. https://itif.org/publications/2021/06/14/current-and-potential-uses-arvr-equity-and-inclusion

Dick's report offers a sweeping overview of how AR/VR technologies can advance equity, empathy, and access, provided that inclusivity is embedded from the start. The document covers assistive applications, bias training, and remote access, while cautioning that immersive technologies are not a panacea for structural inequity. Its analysis of AR/VR as both assistive and transformative aligns closely with my project's mission to reimagine XR as a platform for universal participation, not just accommodation.

Elor, A., & Ward, J. (2021). Accessibility Needs of Extended Reality Hardware: A Mixed Academic-Industry Reflection. Interactions, May–June 2021. http://crossmark.crossref.org/dialog/?doi=10.1145%2F3457877

Elor and Ward reflect on collaborative efforts to assess and advocate for accessibility in XR hardware from both academic and industry perspectives. They emphasize the critical importance of including people with disabilities in hardware design processes from the outset to ensure usability and impact. This article provides my thesis with foundational arguments for participatory design and early stakeholder inclusion, reinforcing the principle that accessibility in XR must be integral rather than retrospective. Their call to action aligns with my project's participatory research and co-design sessions.

Fox, D., Li, A., Pandey, A., & Kar, R. (2019). Augmented Reality for Visually Impaired People (AR for VIPs). University of California, Berkeley, School of Information, Capstone Report.

This capstone report details the development and user testing of an AR application designed to help visually impaired users navigate unfamiliar environments using spatial mapping and machine vision. By incorporating user feedback, the authors demonstrate that context-aware sonification and text recognition can significantly enhance independence. This work directly informs my thesis's advocacy for participatory design and hands-free, multi-sensory interfaces in XR, illustrating successful strategies for user empowerment through AR.

Fox, D. R., Ahmadzada, A., Wang, C. T., Azenkot, S., Chu, M. A., Manduchi, R., & Cooper, E. A. (2023). "Using Augmented Reality to Cue Obstacles for People with Low Vision." Optics Express, 31(4), 6827–6836.

Fox et al. develop and evaluate an AR prototype for obstacle detection, testing visual cueing strategies for users with low vision. Their results demonstrate the effectiveness of world-locked, high-contrast cues compared to heads-up directional indicators, providing empirical guidance for assistive AR design. This research substantiates key technical decisions in my thesis, especially the emphasis on adaptable, context-aware visual feedback to enhance navigation and environmental awareness for users with sensory impairments.

Geerts, D., Vatavu, R.-D., Burova, A., Vinayagamoorthy, V., Mott, M., Crabb, M., & Gerling, K. (2021). Challenges in designing inclusive immersive technologies. In Proceedings of the 20th International Conference on Mobile and Ubiquitous Multimedia (MUM 2021) (pp. 182–185). https://doi.org/10.1145/3490632.3497751

This panel discussion synthesizes the multifaceted barriers to inclusion in immersive technologies like VR and AR, extending the focus beyond disability to situational and cognitive challenges. The authors stress that inclusivity must account for a wide range of user abilities, contexts, and preferences, noting persistent gaps in accessible interaction paradigms and co-design with people with disabilities. This work directly informs my research's inclusive design pillars by highlighting the importance of adaptability, user-centered methodologies, and intersectional accessibility challenges.

Gong, W., Xiao, L., Wang, X., & Lee, C. H. (2020). Dots—An inclusive natural user interfaces (NUI) for spatial computing. In MobileHCI '20 Extended Abstracts (pp. 1–4). https://doi.org/10.1145/3406324.3410715

This paper introduces "Dots," a customizable natural user interface system enabling users with physical disabilities to interact with spatial computing environments. The two-point model underpinning Dots facilitates adaptive interaction patterns through wearable sensors, empowering users to tailor XR experiences to their own body capabilities. Gong et al. demonstrate the importance of flexible, customizable interfaces for achieving genuine inclusivity in spatial computing—a principle central to my thesis.

Hall, E. (2018). Conversational design. A Book Apart.

Erika Hall's treatise on conversational design provides a blueprint for crafting digital interactions that mirror natural human dialogue. By centering the design process on language, intent, and empathy, Hall advocates for systems that are intuitive and genuinely responsive. This book's practical frameworks for voice-driven and conversational interfaces directly inform my thesis's approach to natural language interaction in XR, reinforcing the argument for dialogue as the foundation of accessible digital experiences.

Hamraie, A. (2017). Entangled principles: Crafting a universal design methodology. In Building Access: Universal Design and the Politics of Disability (pp. 223–233). University of Minnesota Press. https://www.jstor.org/stable/10.5749/j.ctt1pwt79d.11

Hamraie critically unpacks the evolution of Universal Design, tracing its shift from disability activism to broader "access-knowledge" and the mainstreaming of universal usability principles. The chapter interrogates both the promise and limitations of the "Principles of Universal Design" and challenges depoliticized approaches that neglect the histories and lived experiences of disability. This critical lens shapes my project's framing of accessibility as both a technical and political mandate for XR.

ANNOTATED BIBLIOGRAPHY

Hoffman, G. (2022). Emotion by design: Creative leadership lessons from a life at Nike. Twelve.

Drawing on nearly three decades at Nike, Hoffman illustrates how creativity and emotional resonance can transform brands and products. The book emphasizes collaborative creativity, bold storytelling, and the pursuit of lasting impact. While not explicitly about accessibility, Hoffman's insights into building human connection through design parallel my project's mission to create XR experiences that are emotionally engaging, memorable, and inclusive at every level.

Jost, M., Luxenburger, A., Knoch, S., & Alexandersson, J. (2022). PARTAS: A Personalizable Augmented Reality Based Task Adaption System for Workers with Cognitive Disabilities. PETRA '22. https://doi.org/10.1145/3529190.3529208

Jost et al. present the PARTAS system, an adaptive, cost-effective AR platform designed specifically for workers with cognitive disabilities in sheltered workshops. The system, using only a camera, projector, and a single-board computer, delivers personalized, contour-based instructions and pick-by-projection functionality, supporting flexible and intuitive task adaptation. Their iterative, user-centered approach involved disabled workers, caregivers, and psychologists to ensure practical usability. This study is a valuable reference for my thesis, highlighting how AR can provide tailored support that bridges the accessibility gap in vocational settings—directly informing my exploration of personalizable XR interfaces for users with varying cognitive abilities.

IDEO.org. (2015). The Field Guide to Human-Centered Design. IDEO.org.

This practical guide outlines the mindsets and methods of human-centered design, emphasizing empathy, iterative prototyping, and creative confidence. The toolkit offers concrete steps for framing challenges, recruiting users, and integrating feedback—practices mirrored in my thesis methodology. IDEO's emphasis on co-design and iterative learning supports the participatory, user-focused approach taken throughout the development and validation of the SpeakEasy XR prototype.

Kee, K. (2022). A simple guide to low muscle tone (hypotonia), diagnosis, treatment and related conditions. Kenneth Kee.

Kee's concise handbook provides foundational knowledge of hypotonia, including diagnostic criteria, clinical presentation, and therapeutic interventions. The book's clear differentiation between muscle tone and muscle weakness, and its focus on daily function and early intervention, ground my thesis's user-needs research and the rationale for voice-first XR interfaces for users with low muscle tone or related neuromotor challenges.

Kolko, J. (2014). Well-designed: How to use empathy to create products people love. Harvard Business Review Press.

Kolko's book champions empathy as the core of effective product development, guiding designers to create solutions that resonate deeply with users' lives and emotions. His actionable process, centered on observation, iteration, and storytelling, has directly influenced my participatory approach to inclusive design in XR. The work's focus on meaningful engagement and emotional impact is central to my thesis's experience design orientation.

Korkiakoski, M., Alavesa, P., & Kostakos, P. (2024). Preference in Voice Commands and Gesture Controls With Hands-Free Augmented Reality With Novel Users. IEEE Pervasive Computing, advance online publication.

Korkiakoski et al. investigate user preferences and usability between voice and gesture controls in hands-free AR (HAR) environments, using the HoloLens 2. Their mixed-method study with novice HAR users found no significant difference in basic usability, but nuanced preferences: gestures were favored for playfulness and intuitiveness, while voice commands were valued for efficiency. Notably, individual variability emerged, with users expressing concerns about learning curves and memorability of commands. This paper is pivotal for my thesis as it reinforces the need to support multiple input modalities in XR—affirming that voice interaction should be context-sensitive and adaptable, especially for accessibility and ease of learning.

Krauß, V., Nebeling, M., Jasche, F., & Boden, A. (2022). Elements of XR prototyping: Characterizing the role and use of prototypes in augmented and virtual reality design. In CHI Conference on Human Factors in Computing Systems (CHI '22). https://doi.org/10.1145/3491102.3517714

This interview-based study examines prototyping practices among industry XR professionals, charting the transition from low- to high-fidelity prototypes and revealing unique challenges in spatial and interactive design. Krauß et al. provide a taxonomy of XR prototypes, highlighting the lack of standardized tools and the need for greater accessibility in prototyping workflows. Their insights inform the iterative, user-centered prototyping approach adopted in my own project.

Lamyman, J. (2024, September 11). Introduction to XR Accessibility. TetraLogical Blog. https://tetralogical.com/blog/2024/09/11/introduction-to-xr-accessibility/

Lamyman's introductory guide distills the practical considerations and evolving standards for XR accessibility. By mapping WCAG guidelines to XR, referencing the W3C XR Accessibility User Requirements (XRAUR), and highlighting the challenges of VR controllers for users with mobility limitations, this article offers actionable insights for designers and developers. The emphasis on offering input choice and enabling personalization directly informs the technical recommendations and user flow design in my thesis project.

Lee, J., Wang, J., Brown, E., Chu, L., Rodriguez, S. S., & Froehlich, J. E. (2024). GazePointAR: A Context-Aware Multimodal Voice Assistant for Pronoun Disambiguation in Wearable Augmented Reality. CHI Conference on Human Factors in Computing Systems (CHI '24). https://doi.org/10.1145/3613904.3642230

This recent study introduces GazePointAR, a wearable AR voice assistant that resolves pronoun ambiguity in user queries by integrating real-time gaze tracking, pointing gestures, and conversation history. By combining these modalities, the system can interpret natural, pronoun-laden speech (e.g., "What is this?") in context, improving accessibility and naturalness in XR environments. The authors' laboratory and diary studies highlight both user enthusiasm for human-like, context-sensitive interaction and the ongoing challenges of implementing robust, privacy-conscious, and explainable multimodal AI. GazePointAR directly informs my thesis's ambition to build more intuitive, accessible voice interfaces in XR, especially for users relying on gesture or gaze over traditional input.

Li, Z., Connell, S., Dannels, W., & Peiris, R. (2022). SoundVizVR: Sound Indicators for Accessible Sounds in Virtual Reality for Deaf or Hard-of-Hearing Users. ASSETS '22. https://doi.org/10.1145/3517428.3544817

SoundVizVR introduces a suite of visual sound indicators—combining mini-maps, on-object signals, text, and icons—to convey auditory information for Deaf and hard-of-hearing VR users. User studies found that combining full mini-maps with on-object indicators best supported sound localization, while additional text and icon cues improved understanding. This research underpins my thesis argument that multimodal feedback is crucial in XR accessibility, particularly for users with sensory impairments, and supports the integration of layered visual sound cues in accessible design.

Li, J. (2024). Beyond sight: Enhancing augmented reality interactivity with audio-based and non-visual interfaces. Applied Sciences, 14(11), 4881. https://doi.org/10.3390/app14114881

Li's article addresses a fundamental gap in AR design: the heavy reliance on visual interfaces that marginalize users with visual impairments. Through developing and evaluating audio-based, non-visual AR prototypes, this research demonstrates that auditory cues can significantly enhance both spatial awareness and usability, benefiting a wider range of users. The findings advocate for integrating audio-based interaction as a core accessibility strategy in AR, underscoring the need to move beyond visual-centric paradigms in immersive technology design. This work supports my thesis's emphasis on multi-sensory, inclusive interaction frameworks.

Madhok, S. S., & Shabbir, N. (2022). Hypotonia. In StatPearls (NCBI Bookshelf). https://www.ncbi.nlm.nih.gov/books/NBK562209/

This medical overview distills the clinical presentation, diagnosis, and management of hypotonia, emphasizing its multifactorial causes (central, peripheral, genetic, metabolic). The article details assessment maneuvers—vertical and horizontal suspension, scarf sign, pull-to-sit—and their relevance to distinguishing hypotonia from weakness. The discussion of multidisciplinary management, including rehabilitation, highlights the complexity and need for adaptive support. For my research, this source grounds the user-needs framing and helps specify the functional requirements for XR accessibility aimed at users with low muscle tone.

Maeda, J. (2019). How to speak machine: Computational thinking for the rest of us. Portfolio/Penguin.

John Maeda demystifies computational thinking and artificial intelligence, offering designers and leaders a high-level vocabulary and conceptual toolkit for collaborating with intelligent systems. Through accessible analogies and practical guidance, Maeda addresses the challenges and ethical questions posed by machine learning and automation. His call for inclusive, transparent, and responsible AI directly echoes my thesis's advocacy for human-centered, intelligible AI in XR environments.

Maran, P. L., Daniëls, R., & Slegers, K. (2022). "The Use of Extended Reality (XR) for People with Moderate to Severe Intellectual Disabilities: A Scoping Review." Technology and Disability, 34, 53–67.

This scoping review surveys the emerging literature on XR-based interventions for people with moderate to severe intellectual disabilities (ID). The authors identify a range of applications—primarily focused on daily living, navigation, and academic skills—but note a lack of large-scale, controlled studies. Importantly, the review emphasizes the need for careful adaptation of XR environments to user needs, considering support structures and implementation factors. The study's findings validate my thesis's focus on adaptable, voice-driven experiences as a pathway for genuine inclusion, and highlight gaps that future work, like SpeakEasy, can address.

Mathew, R., Mak, B., & Dannels, W. (2022). Access on Demand: Real-time, Multi-modal Accessibility for the Deaf and Hard-of-Hearing based on Augmented Reality. In Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22), 6 pages. https://doi.org/10.1145/3517428.3551352

This experience report evaluates Access on Demand (AoD), an AR-based application providing real-time captioning and ASL interpretation for Deaf and Hard-of-Hearing users. The study foregrounds firsthand user experiences with AR smart glasses, highlighting both the benefits and current technological limitations. AoD's multi-modal approach exemplifies the potential for XR to deliver on-demand accessibility services beyond traditional environments. This work supports my thesis by evidencing the practical value and user-centered considerations of real-time adaptive accessibility, underscoring the necessity of including DHH voices in XR design and evaluation.

Michalski, S. C., Szpak, A., Ellison, C., Cornish, R., & Loetscher, T. (2022). "Using Virtual Reality to Improve Classroom Behavior in People with Down Syndrome: Within-Subjects Experimental Design." JMIR Serious Games, 10(2), e34373.

This study examines the impact of VR experiences on classroom behavior in young adults with Down syndrome, finding significant improvements in mood, attention, and engagement following both VR and conventional drawing activities. The paper also discusses the feasibility of VR for this population, acknowledging challenges such as vision impairment and cybersickness. The findings support the potential of immersive technologies to foster positive behavioral outcomes, aligning with my thesis's focus on inclusive, motivating XR activities for neurodiverse users.

ANNOTATED BIBLIOGRAPHY

Milliken, Neil. "Neurodiversity Celebration Week." CAW Blog, 21 March 2022. https://blog.caw.ac.uk/index.php/caw-news/neurodiversity-celebration-week/.

Neil Milliken's advocacy for neurodiversity is a clarion call to embrace cognitive differences as a societal asset. This quote, drawn from an article celebrating Neurodiversity Celebration Week, is particularly poignant for my thesis, which seeks to harness the unique perspectives and abilities of neurodiverse individuals in enhancing XR accessibility. Milliken's words reinforce the thesis's underlying premise that diversity enhances creativity and innovation within design. His call to celebrate neurodiversity aligns with the project's commitment to create XR environments that are not only accessible but also diverse and vibrant, reflecting the rich tapestry of human cognition.

Mogavi, R. H., Hoffman, J., Deng, C., Du, Y., Haq, E.-U., & Hui, P. (2023). Envisioning an inclusive metaverse: Student perspectives on accessible and empowering metaverse-enabled learning. In Proceedings of the Tenth ACM Conference on Learning @ Scale (L@S '23), 346–355. https://doi.org/10.1145/3573051.3596185

Mogavi et al. present qualitative findings from interviews with disabled students about their aspirations and concerns for the metaverse in higher education. Their REEPS framework—Recognition, Empowerment, Engagement, Privacy, Safety—articulates the core values that should guide inclusive metaverse development. This user-driven perspective underlines the necessity of early and ongoing engagement with diverse users, aligning directly with the participatory design principles in my research.

Mott, M., Cutrell, E., Gonzalez Franco, M., Holz, C., Ofek, E., Stoakley, R., & Morris, M. R. (2019). Accessible by Design: An Opportunity for Virtual Reality. ISMAR 2019 Workshop on Mixed Reality and Accessibility.

This position paper contends that accessibility in VR is too often an afterthought and argues for its integration from the earliest stages of development. The authors outline five key areas for accessible VR—including content, interaction techniques, and hardware—urging the field to develop cross-industry standards and guidelines. This work is vital to my thesis as it situates accessibility as a universal design opportunity rather than a constraint, highlighting actionable areas for improvement and the legal, technical, and ethical imperatives for inclusion

Moustafa, R. S., Karhu, H., Andberg, S., & Bednarik, R. (2023). Seeing Through Their Eyes - A Customizable Gaze-Contingent Simulation of Impaired Vision and Other Eye Conditions Using VR/XR Technology. ETRA '23. https://doi.org/10.1145/3588015.3590110

This work describes a customizable VR/XR system that simulates impaired vision conditions using gaze-contingent rendering with a Varjo Aero headset and eye tracking. The approach enables realistic, first-person simulation of visual impairments for diagnosis, training, and empathy building, offering adjustable artifacts and data recording. For my thesis, this paper highlights the importance of simulating user perspectives to foster empathy and improve the design of accessible XR environments, providing a framework for user-centered, disability-aware development.

Mozilla (2020). Creating accessible VR experiences. Mozilla Hubs VR Accessibility Guidelines. https://hubs.mozilla.com

Mozilla's guidelines showcase pragmatic steps for embedding accessibility into social VR environments, particularly through the lens of the Hubs platform. By adopting browser-level accessibility features and prioritizing user-friendly interfaces—such as WCAG-compliant color palettes and minimal account barriers—this resource grounds abstract accessibility principles in actionable design tactics. Its emphasis on iterative, community-informed design and cross-sensory accommodations is highly relevant to the practical modules of my prototype and the process book.

Norman, D. A. (2023). Design for a better world: Meaningful, sustainable, humanity centered. The MIT Press.

In this manifesto, Norman urges designers to prioritize sustainability, human well-being, and global impact, advocating for systemic change in how products and environments are conceived. He proposes frameworks for meaning-driven, humanity-centered design—relevant to my thesis's aim to move beyond technical accessibility toward experiences that foster connection, equity, and long-term value in XR.

Norman, D. A. (2004). Emotional design: Why we love (or hate) everyday things. Basic Books.

Norman's exploration of emotional responses to everyday objects introduces a tripartite model—visceral, behavioral, and reflective design—that explains why some products are beloved and others frustrating. His work underscores the importance of emotional engagement and narrative in design, reinforcing my project's dual emphasis on usability and delight in accessible XR systems.

Norman, D. A. (2013). The design of everyday things: Revised and expanded edition. Basic Books.

A seminal text in design theory, Norman's book dissects why users struggle with poorly designed products and champions affordances, feedback, and signifiers as keys to intuitive design. The revised edition brings the classic principles into the context of modern technology, supporting my thesis's commitment to human-centered, error-tolerant, and inclusive XR interaction models.

Northover, Kylie. "Oliver Sacks: His Best Quotes." The Sydney Morning Herald, August 31, 2015. https://www.smh.com.au/entertainment/books/oliver-sacks-his-best-quotes-20150831-gjbdkz.html.

In this collection of profound musings, Oliver Sacks inspires a shift in perception towards disability, advocating for the recognition of unique attributes as powerful assets. This quote is integral to my thesis, as it champions the notion that inclusive design in XR should harness and elevate the diverse capabilities of individuals with disabilities. It influences my research direction, underlining the potential within the XR domain to create environments where every user's attributes, including those stemming from disabilities, are not just accommodated but valued as key contributors to the richness of the human experience. Sacks' vision propels the thesis's aim to develop XR technologies that are not merely accessible but celebratory of human diversity, fostering a digital realm where each person's uniqueness is acknowledged as a strength.

Nussbaumer, L. L. (2019). Inclusive Design: A Universal Need. South Dakota State University.

Nussbaumer's book serves as a foundational text on the evolution, principles, and application of inclusive design, extending the conversation beyond accessibility to address the diversity of users' abilities and contexts. The discussion of universal and inclusive design philosophies, case studies, and practical guidance for integrating inclusivity throughout the design process informs my thesis's theoretical grounding. By advocating for environments and products that anticipate difference rather than retrofit for it, the work validates my focus on proactive, voice-driven XR experiences.

O'Connor, J., Sajka, J., White, J., Hollier, S., & Cooper, M. (2021). "XR Accessibility User Requirements." W3C Working Group Note. https://www.w3.org/TR/xaur/

This W3C document systematically outlines user needs and requirements for accessibility in XR, providing a comprehensive overview of challenges across sensory, physical, cognitive, and neurological domains. It articulates the necessity for multi-modal input, customizable controls, and adaptable output modalities, highlighting areas such as motion-agnostic interactions and the provision of critical messaging in immersive environments. This resource serves as an authoritative reference for my thesis, grounding my design framework in international accessibility standards and informing the inclusive intent behind the SpeakEasy prototype.

Pangilinan, E., Lukas, S., & Mohan, V. (2019). Creating augmented and virtual realities: Theory and practice for next-generation spatial computing. O'Reilly Media.

This practical guide covers the full pipeline of AR and VR development, blending foundational theory with hands-on approaches to design, prototyping, and cross-platform engineering. Its emphasis on human-centered interaction, content creation, and real-world use cases aligns closely with my thesis's method of iterative, experience-driven XR development, making it a valuable resource for both theory and practice.

Papanek, V. (2005). Design for the real world: Human ecology and social change (2nd ed.). Chicago Review Press.

Papanek's influential critique of mainstream design calls for social responsibility, ecological awareness, and the prioritization of genuine human needs. He challenges designers to address global problems with solutions that are sustainable and equitable. This perspective underpins my thesis's ethos of inclusive, universally beneficial XR systems.

Parker, C., Yoo, S., Lee, Y., Fredericks, J., Dey, A., Cho, Y., & Billinghurst, M. (2023). Towards an Inclusive and Accessible Metaverse. CHI EA '23. https://doi.org/10.1145/3544549.3573811

This workshop report summarizes efforts to define accessibility and inclusion requirements for the emerging Metaverse, emphasizing the risk of fragmentation and exclusion if standards and best practices are not developed. Key themes include the need for adaptive hardware, evaluation methodologies, and participatory design approaches. The paper supports my thesis in its argument that inclusivity must be intentionally designed into XR ecosystems from the outset, highlighting the importance of co-creation and policy alignment.

Pearl, C. (2016). Designing voice user interfaces: Principles of conversational experiences. O'Reilly Media.

Pearl's definitive guide to VUI design explores the principles of creating effective, natural, and error-tolerant voice interfaces. The book provides concrete strategies for dialog management, feedback, and accessibility—knowledge directly applied in my thesis's development of voice-driven, accessible XR applications.

René, G., & Mapes, D. (2019). The Spatial Web: How Web 3.0 will connect humans, machines, and AI to transform the world. BookBaby.

René and Mapes theorize the Spatial Web—a convergence of XR, AI, IoT, and blockchain—arguing for new standards and ethical frameworks as the digital and physical worlds merge. Their vision supports my thesis's long-term roadmap for accessible, interoperable, and value-driven XR platforms that prioritize sustainability and inclusivity.

Rob Scott. (2022, January 27). What Is Speech Recognition Technology in VR? XR Today. https://www.xrtoday.com/virtual-reality/what-is-speech-recognition-technology-in-vr/

Scott provides an accessible overview of the current state of speech recognition in VR, emphasizing its growing importance for hands-free, intuitive navigation, and digital agent interaction. The article details technical limitations (accuracy, ambient noise, accent variability) and the emergence of SDKs (e.g., Meta Voice SDK) that lower barriers to integration. This source is relevant for my thesis as it articulates both the promise and ongoing challenges of voice-driven XR interaction, affirming the need for robust error handling and fallback strategies in voice-based accessible systems.

Ruh, Debra. "The only disability is when people cannot see human potential." Morning Lazziness. https://www.morninglazziness.com/quotes/accessibility-quotes/.

Debra Ruh's quote is a poignant reminder of the importance of perceiving abilities in place of disabilities. This statement is a critical reflection on society's often limited view of individuals with disabilities, suggesting that the real limitation is not in the individuals' conditions, but in society's failure to recognize their potential. This perspective is a driving force behind my thesis on XR accessibility, as it informs the approach of looking beyond the conventional understanding of disabilities, advocating for a design philosophy that uncovers and leverages the unique strengths and capabilities of every individual. Ruh's insight challenges us to build XR environments that are not only accessible but also nurturing of the potential in all users .

Rubin, P. (2018). Future presence: How virtual reality is changing human connection, intimacy, and the limits of ordinary life. HarperOne.

Rubin investigates VR's potential to fundamentally reshape human relationships and connection, highlighting both its capacity to induce empathy and the ethical complexities it introduces. His narrative-style exploration of VR's social and emotional impact enriches my thesis's framing of XR as a tool for building not just accessible, but also meaningful and emotionally resonant experiences.

Ryskeldiev, B., Ochiai, Y., Kusano, K., Li, J., Saraiji, Y., Kunze, K., Billinghurst, M., Nanayakkara, S., Sugano, Y., & Honda, T. (2021). Immersive Inclusivity at CHI: Design and Creation of Inclusive User Interactions Through Immersive Media. CHI Conference Extended Abstracts, 4 pages. https://doi.org/10.1145/3411763.3441322

This workshop paper articulates the convergence of immersive media (AR/VR/XR) and inclusive design, calling for the creation of truly accessible multimodal environments. By surveying the current landscape and proposing a research and evaluation agenda, the authors argue for methods that foreground users with diverse abilities—whether permanent, temporary, or situational. Their examples and taxonomy of immersive inclusive design reinforce the need for iterative, user-driven innovation. This aligns closely with my thesis approach and situates my work within a wider research community prioritizing XR accessibility.

Schmelter, T., Rings, S., Kruse, L., Steinicke, F., Karaosmanoglu, S., & Hildebrand, K. (2023). Towards More Inclusive and Accessible Virtual Reality: Conducting Large-scale Studies in the Wild. CHI EA '23. https://doi.org/10.1145/3544549.3583888

This paper introduces the VITALAB.mobile, a mobile, wheelchair-accessible VR lab enabling inclusive research and therapy studies in diverse, real-world contexts. The mobile lab brings VR access to users often excluded due to mobility or location barriers, supporting a wider range of participants and applications, including rehabilitation and exergames. The concept reinforces my thesis's claim that accessibility in XR requires adaptable, context-aware solutions, and that research infrastructure itself must evolve to reach underrepresented users.

Shneiderman, B. (2022). Human-centered AI. Oxford University Press.

Shneiderman's optimistic but practical roadmap for AI advocates for systems that augment human abilities, preserve dignity, and foster trust. The book's fusion of ethics, usability, and technical design principles echoes my thesis's insistence on accessible, transparent, and empowering AI within voice-driven XR environments.

Simon-Liedtke, J. T., & Baraas, R. (2022). Towards eXtended Universal Design XR in Education. In I. Garofolo et al. (Eds.), Transforming our World through Universal Design for Human Development. IOS Press. https://doi.org/10.3233/SHTI220865

Simon-Liedtke and Baraas report on barriers to XR adoption in education due to lack of universal design, based on focus groups with educators and disability advocates in Norway. They outline opportunities, such as supporting students with Down syndrome and ADHD, but also significant challenges in hardware fit, multimodality, and interface accessibility. The study calls for standards, co-creation, and funding for inclusive XR. This aligns with my thesis's advocacy for universal, not just targeted, accessibility measures in XR learning environments.

Wu, D. (2024). Spatial design: Breaking the 2D paradigm. XReality Pro.

Wu's book tackles the challenges of transitioning from 2D to spatial (3D) design, offering a set of practical principles, case studies, and industry interviews. The emphasis on visibility, natural user interfaces, and broad-based skills is directly relevant to my thesis's goal of making spatial, voice-first XR experiences intuitive and accessible for a wide range of users.

ANNOTATED BIBLIOGRAPHY

Xie, J., & Zhao, T. (2022). “VR Scene Taxonomy: Designing Accessible Scene Viewing Techniques.” In Proceedings of the ACM Symposium on Virtual Reality Software and Technology.

This paper introduces a taxonomy of scene viewing techniques in VR, specifically considering the accessibility needs of users with disabilities. The authors propose a framework to evaluate and design VR scene viewing modes for inclusivity, accounting for challenges faced by users with low vision, motor impairments, or cognitive differences. For my thesis, this work provides both a conceptual scaffold and actionable design criteria for ensuring the core navigation and orientation tasks in immersive XR are not only technically possible but also genuinely usable for diverse populations. The taxonomy's approach helps validate the iterative testing and adaptive UI strategies in my own prototype.

YourTechDiet. (2025). “Voice Assistants: Future Trends Beyond 2025.” https://yourtechdiet.com/blogs/future-of-voice-assistants-a-look-beyond-2025/

This industry analysis projects major advancements for voice assistants in the coming decade, including improved natural language understanding, proactive contextual support, and integration with AR devices. The article predicts a shift from reactive commands to anticipatory, agentic interaction models, emphasizing the growing role of voice as a universal interface. These trends contextualize my thesis's technological roadmap, reinforcing the strategic importance of voice-first design for accessible, multimodal XR experiences.

Zallio, M., & Clarkson, P. J. (2022). Designing the metaverse: A study on inclusion, diversity, equity, accessibility and safety for digital immersive environments. Telematics and Informatics, 75, 101909. https://doi.org/10.1016/j.tele.2022.101909

Through expert interviews, this study foregrounds the ethical, social, and practical imperatives for designing an inclusive metaverse. Zallio and Clarkson identify best practices and persistent gaps across inclusion, accessibility, and digital safety, offering actionable recommendations for industry and research. Their findings reinforce the need for intersectional design strategies and iterative assessment, echoing the structure of my thesis's framework for inclusive XR experiences.

Zhao, J., Parry, C. J., dos Anjos, R., Anslow, C., & Rhee, T. (2020). “Voice Interaction for Augmented Reality Navigation Interfaces with Natural Language Understanding.” In IEEE International Symposium on Mixed and Augmented Reality (ISMAR).

This paper evaluates the effectiveness of natural language understanding (NLU) in AR navigation interfaces, introducing the VOARLA system and measuring user accuracy, productivity, and learning curve. Findings show that NLU increases command accuracy and reduces user cognitive load, though gains in productivity are not guaranteed. For my thesis, this research validates the integration of NLU in voice-driven XR, supporting the premise that semantic flexibility and intuitive phrasing are essential for inclusive, hands-free interaction models.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support, guidance, and encouragement of many remarkable individuals and communities.

I am deeply grateful to my committee chair, Dr. Yoon Chung Han, for her expert advice, high standards, and steady mentorship throughout every stage of this project. Thank you to Prof. Marjan Khatibi and Dylan Fox for their invaluable insights, encouragement, and thoughtful critique. Dylan's ongoing work in XR accessibility and his willingness to share lived experience shaped both my research and my perspective.

Special thanks to the members of the XR Access community for sharing your stories, challenges, and vision for a more inclusive future. Your feedback and advocacy inspired many of the decisions behind SpeakEasy.

To my family—my son Mason, whose curiosity and resilience inspire everything I do; to Cherene and her children, for their patience, love, and understanding during late nights and long sprints; and to my friends and cohort at San José State University for pushing me to think deeper and keep going.

I would especially like to thank my parents, whose unwavering support, encouragement, and belief in my path made this journey possible. Your love and guidance have shaped not just this thesis, but every creative pursuit I have undertaken.

Thank you as well to the participants, user testers, and advocates who volunteered their time, energy, and honest feedback. Your lived experiences and willingness to experiment made this project real.

Finally, gratitude to the communities, toolmakers, and open-source contributors whose work built the foundations for this thesis—especially those developing accessible XR, voice AI, and inclusive design tools. I hope this work pays your vision forward.

ABOUT ME

With a background in graphic design and engineering and a deep interest in user interaction, I wanted to focus on Experience Design to blend technical expertise with practical, impactful user experiences.

My past experience has been working in Extended Reality (XR), striving to make these digital environments universally navigable. By tackling the shortcomings in existing XR interfaces, I am committed to designing systems that are not just usable, but anticipatory and responsive to a spectrum of user needs.

My aspiration is to alter the trajectory of design towards practices that are genuinely inclusive and functionally relevant, using technology to enhance, rather than complicate, human connections.

A SJSU M.Des Thesis Project

Updated as of May 2025.