Rap is Art: Creating Spaces for the Interpretation of Misunderstood Voices

TeAiris Majors
Georgia Institute of Technology
tmajors6@gatech.edu

ABSTRACT

This paper explores the integration of rap music, an auditory art form conveying powerful sociopolitical commentary, into visual arts spaces to create immersive multimedia experiences. A virtual museum was developed in Unity, spatializing music from prominent rap artists to amplify their messages. Interactions allow users to visualize lyrics and access sources linked to deeper analysis. Parallels are drawn to soundscape composition techniques and the World Soundscape Project's exploration of acoustic ecology. Goals include highlighting lyrical depth in rap by extracting and projecting vocal and song lyrics, uniting visual and auditory arts through an inclusive museum space, and providing an innovative template for future virtual experiences. This paper argues that rap music deserves recognition as impactful artistic expression, not just mere entertainment, containing deeper meaning beyond its aural perspective. It concludes that embracing rap music in cultural institutions and virtual interactive experiences can increase accessibility and awareness, foster dialogue about social justice, identity, and empowerment, and resonate with more diverse audiences.

1. INTRODUCTION

Paintings and music both serve as channels for artistic expression [1], evoking emotions and opportunities for personal perceptions to emerge. One being a visual medium and the other an auditory, each having its own unique way of engaging and impacting our surroundings. Museums are great spaces for centralizing art, as they capture the history, diverse cultures, and stories of a particular time or space. Traditionally known for showcasing visual art, museums have also explored embracing the history and significance of music genres such as hip-hop culture and rap music by means of photos, visuals, relevant artifacts, and auditory experiences [2]

Rap music has embraced various forms of storytelling and narrative techniques, allowing artists to provide a voice for marginalized communities and shed light on their experiences [3]. Unfortunately, stereotypical narratives and misconceptions about rap music focus solely on explicit language, violence, and misogyny, failing to recognize the deeper messages and social commentary embedded within the lyrics [4, 5]

Fortunately, artists within the genre continue to push boundaries and challenge these stereotypes, using their platform to address important social issues such as systemic racism, poverty, police brutality, and inequality [6]. Artists, such as Kanye West, have used their platform to challenge societal norms and initiate meaningful discussions [7]. Kendrick Lamar's album "To Pimp a Butterfly" is a powerful example of the social and political commentary that rap music can provide, addressing issues of racial inequality, self-love, and the Black experience in America. Undoubtedly, the impact of rap music in society goes beyond mere entertainment.

The discourse of rap music cannot all be fully understood without contextualizing it within the environment in which it emerges. Some may turn to the news to stay informed about current events. However, this is usually derived from an outside perspective, often influenced by dominant narratives and biased interpretations. While this can provide some level of awareness, rap music offers a unique and unfiltered perspective within communities.

Similarly, soundwalks and soundscape compositions provide a different way of experiencing and understanding our environment [8]. By incorporating the sounds of everyday life, these audio experiences allow us to listen to our surroundings with intention and curiosity. The composers of this genre target the nuances of the sounds around us, highlighting the intersection between sound, space, and emotion.

With the advances of technology, now it is even easier to deliver art and entertainment to a wider audience through the use of virtual experiences [9]. Creators have access to tools, such as game engines accompanied with spatial audio features, to design the virtual spaces of our imagination, inviting remote elements that may have not co-existed or shared similar physical and/or aural environments.

This project aims to explore this concept of coexistence and augment the perspective in which we experience art and music. By localizing the music, lyrics, and internal voices of prominent rap artists in a virtual space, I aim to amplify these narratives and foster dialogue about social justice, identity, and empowerment.

2. RELATED WORK

Composers such as Hildegard Westerkamp have really pushed to expose the deeper embedded meanings of sounds in our environment and the roles they have in society [10]. Part of this force is to deliver an experience of a space that we may have not traveled to. Another part, being the soundwalk, is to provide a different perspective in real time. This project, while not an exact mirror of their philosophies, aims to explore innovative ways to experience sound in an environ-

ment beyond static recordings. By making a system that is somewhat varying, this soundwalk experience invites multiple perspectives within space.

Soundscape composers traditionally captured sounds in a space as they present themselves in nature, later reproducing the sound as a form of art. Some composers in this genre even use special techniques to process the source, yet still resemble its native origin. With advances in technology, such as 3d renders, we have access to defining these virtual spaces. In this project, I make use of a virtual museum. While a museum does not have a distinct sound itself, the augmentation of its use echoes these concepts.

Observing the work of Annea Lockwood, mother nature has voice that can speak for itself. Still, we may stumble upon more of her universal beauty if we search deeper. Advances in audio recording technology have given us innovative ways to capture and reproduce the sounds we hear in our daily and remote physical environments. For example, a listening session and talk with artists, such as Listening to Rivers - with Leah Barclay and Annea Lockwood include examples of hydrophones (waterproof omnidirectional microphones) used to capture the auditory scene beneath earth's oceans and rivers [11] Likewise, in this project, I aim to expose a messages that were always present, but submerged beneath instrumental beats. I also strive to express the many ways in which sound can be experienced and appreciated in different environments.

Stepping back to its roots, the World Soundscape Project (WSP) was founded by R. Murray Schafer at Simon Fraser University in the early to study acoustic environments [12]. In relation, soundscape composition composers, such as Hildegard Westerkamp aimed to emphasize and explore the interrelationship between sound, nature and society (Acoustic ecology) [10]. In her article "Linking soundscape composition and acoustic ecology ", Westerkamp discusses the composer's affordances to convey a point of view through soundscape pieces and their impact [10]. While flexible in terms of definition, soundscape compositions are not all mere reproductions of a recorded environment. In Westerkamp's Beneath the Forest Floor, she explores transformations of sounds within an environment to metaphorically represent sounds akin to a typical rainforest [13]. She also uses unconventional ways to capture sound in the environment in addition to stationary field recordings. Barry Truax's analysis of this practice adequately justifies the flexibility within the soundscape genre [14]. This practice can synonymous to the art of sampling in hip-hop music. While a sample may undergo many chops, manipulations, and processing, it can always be traced back to its origin.

3. GOALS

One focal goal of this project is to bring attention to the powerful messages and social commentary embedded within rap music. Although we normally listen to songs as they are produced and delivered, I believe that by extracting vocal stems, the interpretation of lyrics can be amplified. I also aim to unite the visual and aural art-forms by creating a space that welcomes them equally. Another goal of this project is to explore the use of spatial audio to realize creative interactions with sound that do not exist in the physical world. Additionally, this project serves as a resource and template for creating virtual soundscapes, and other multimedia interactive experiences.

4. IMPLEMENTATION

4.1 Development Environment

Unity Game Engine was used as the main development platform for this project. With features such as sound spatialization, 3d model importing, and video playback, it was an ideal choice. Another advantage of using Unity is its well supported developer community and the Unity Asset Store. This allows for easy integration of extended features and assistive editor add-ons. Furthermore, the access to visual scripting tools, such as Flow Canvas[15] [1] empower artists, even with little programming knowledge, to implement innovative interactions and behaviors and allowing for more experimentation and creativity in the development process.

4.2 Visual Design

To emphasize music as an art form, the virtual space was designed to resemble a small museum. Rap music can sometimes be regarded to as just entertainment or "noise", even if regarded in a metaphorical sense, rather than a true form of art [16] . The walls are decorated with styled paintings of rap artists to emphasize their status as worthy works of art themselves. In addition, a few of the paintings display famous rap lyrics of songs that have had a significant impact in the rap genre. The canvases were intentionally designed to be very reflective. Since the lighting was baked (rendered) in advance, the artwork is still reflected even when changing textures for *Mode 2* discussed later in this paper. The walls themselves were programmed to be forever changing. This is implemented to resemble the evolution and growth that rap music and hip-hop culture has seen over the past 50 years since its birth in 1973. It is also meant to serve as a symbol of inclusiveness and welcome diversity within this artistic space. The color of the walls is determined by the hue value of its inherited material; see Figure 2. Over a period of 30 seconds, the hue is increased from 0 to 360 to complete the color wheel. This method is triggered by a custom Update function, which has an interval of 30 seconds also. Sharing space with the canvases, styled text accompanies the walls. The Kayne.rest API is used to retrieve quotes made by the rapper Kayne West, in real-time; see Figure 1. This is included to further support the project's goals and amplify voices that are often suppressed and/or silenced.

4.3 Song Selection

The selection of songs for this project were made with careful consideration to represent a diverse range of rap artists and themes. The songs included are as follows:

"Neighbors" by J. Cole: portrays his personal experiences with racial profiling and the assumptions made about

[1] Flow Canvas "a powerful and feature-rich Visual Scripting Solution for Unity, empowers others to create and manipulate virtually any aspect of game-play elements for your games in a very similar fashion to Unreal Blueprints, but with far less programming knowledge required all around".

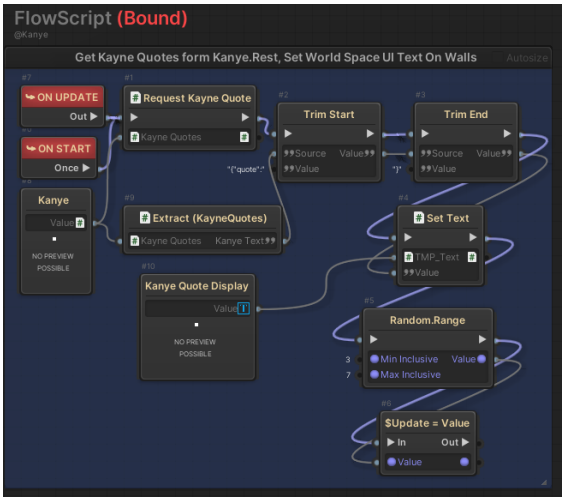

Figure 1. This Flow Script Graph contains nodes to request and display quotes created by Kayne West, delivered from https://api.kanye.rest.

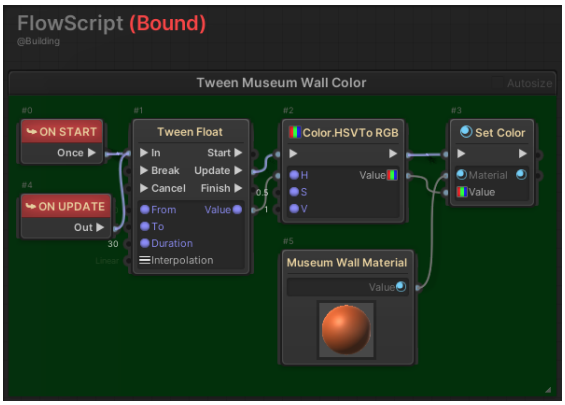

Figure 2. This image displays logic created using Flow Canvas, a third-party visual scripting solution for Unity Game Engine. This Flow Script Graph contains nodes to implement changing the hue of walls of the environment at a defined interval of 30 seconds, in an infinite loop.

him as a successful African American man living in an affluent neighborhood.

"Apparently" by J. Cole: reflects on his life and his relationships, particularly with his mother. This song carries themes of repentance, forgiveness, and understanding the significance of loved ones.

"Alright" by Kendrick Lamar: viewed as an anthem for hope and resilience. It delves into the struggles faced by the African American community and emphasizes a message of determination and confidence that things will get better.

"DNA" by Kendrick Lamar: showcases Kendrick Lamar's complex lyrical ability and strong pride in his African American heritage. This song explores his personal and cultural identity and the qualities that define him.

"All Falls Down" by Kanye West: critiques consumerism, self-consciousness, and the search for self-worth in material possessions. It also discusses the challenges faced by African American communities and the pressure to conform to societal expectations.

"U.N.I.T.Y." by Queen Latifah: a powerful feminist anthem that confronts issues of sexism, misogyny, and disrespect towards women. This song advocates for female

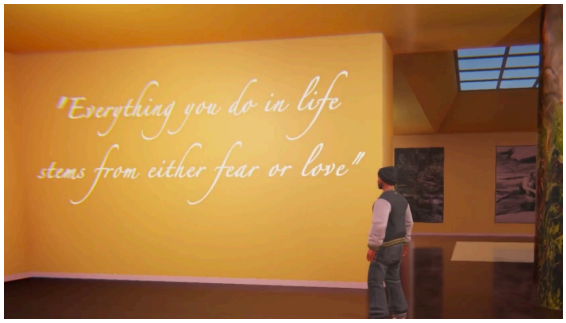

Figure 3. This image displays a wall decorated with text delivered from https://api.kanye.rest.

empowerment and respect for women from all walks of life.

Individual audio files were essential for the purpose of this project in order to facilitate analysis and exploration of the lyrics, stem separation, and other exploratory audio processing. Each song used in this project was purchased from Apple's iTunes Store. A digital copy was then downloaded locally. These files are used specifically for research and educational purposes and within the boundaries of fair use. To comply with copyright laws, the audio files will not be shared or distributed with future template versions of this project.

4.4 Interaction and Sound Design

4.4.1 Mode 1

The interaction and sound design for this project were crucial in creating an immersive and engaging experience [9]. To organize the experience, two modes were implemented. In *Mode 1*, I included a human character for realism, which a user can navigate through the virtual museum. Virtual pillars are placed in the space containing videos of hip-hop pioneers such as Afrika Bambaataa and Grandmaster Flash, symbolizing the foundation of the rap genre. As a user approaches each pillar, the video's sound is spatialized, accompanied by a visual display of the song's lyrics projected into the virtual environment. As the user gets closer, the acapella version of the song is tweened in.

To achieve lyrical text projection, the third-party asset "sound text transcriber asset" was used. This asset uses OpenAI's automatic speech recognition system "Whisper" to transcribe audio, then divides the text into synced phrases [17]. Next, the synced text events are passed into a text mesh pro component. Lastly, another third-party asset, "3d text asset," was used to instantiate and animate 3d text objects into the environment, while fading after a specified duration. This integration aims to provide a multi-sensory experience [18], allowing the users to engage with both the auditory and visual elements. Beyond the pillars, users can enjoy viewing the art work just as you would in a museum.

4.4.2 Mode 2

In *Mode 2*, I aimed to emphasize the goal of highlighting lyrical depth and intricacy in rap music. The character is now modeled as a robot. This interacts with ideas presented in Hildegard Westerkamp's article "Linking Soundscape Composition and Acoustic Ecology". She proposed

Figure 4. This image displays the canvases on the wall of the museum. It also includes the 3d world space text used to display lyric visuals with spatialized audio.

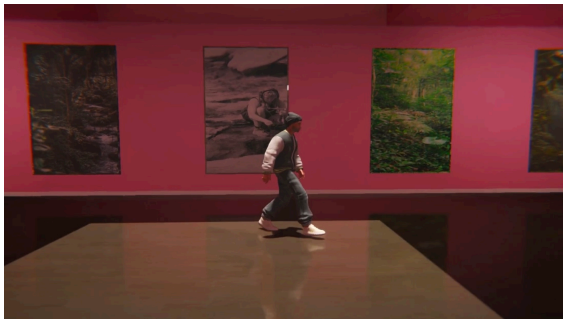

Figure 5. This image displays the area created as a tribute to Soundscape Composition composer Hildegard Westerkamp.

that the human ear has the ability to focus in and out, but the microphone's way of hearing is non-selective[10]. The robot character here serves as a neutral observer, encouraging the listener to listen from a similar perspective.

Triggered using the bottom face button of a controller, or the space bar, most of the art work in the virtual museum transforms into live music videos of the selected rap songs by J. Cole, Kanye West, Kendrick Lamar, and Queen Latifah projected from each canvas; see Figure 4 All videos references were stored in a list, then randomly distributed among the canvases, ensuring that each experience is unique. The Fisher-Yates shuffle algorithm was used in a C# script to achieve this function. The playback start times of each video are randomized also. To achieve this, the Get Frame Count method is used on each Video Component. Next, a random range function is called using 0 as the minimum, and the total frame count for the video clip as the maximum. To set the playback, the Set Fame method is called using the randomly generated number as an argument.

Again, the sound of each video is spatialized to enable a controlled listening environment. Although the videos are positioned closely for an intimate feel, the music occupies its own local space. The maximum distance variable for the audio source spatialization was purposely set to slightly overlap with adjacent videos. This decision was made to explore the artists as two unique voices sharing the same resonate space to convey similar messages from different perspectives. Using a similar function in Mode 1's pillars, a box trigger was used to detect the presence of the user near a video, calling the Acapella function which tweens the transition from the full audio track to the isolated vocal stem.

4.5 Soundscape Composition Tribute

While conceptualizing and fulfilling the creation of this experience, I was greatly inspired by the history, purpose, and practice of soundscape composition and composers of the genre. To express my gratitude and appreciation for their efforts, a portion of the space was dedicated one particular composer, Hildegard Westerkamp, to display one of her aural artworks "Beneath the Rain Forest" [13]; see Figure 7. A pillar, accompanied by a video recorded in a rain forest, is also placed near the entrance of this area to situate users before entering. While placing Westerkamp's voice among the other rap artists was considered, I chose to separate the two to clearly express the intended gesture.

4.6 Song Comparison

As an extension of the first two modes, I have added a few prototype examples to further enrich this virtual experience. To start, I created a matrix in Word to cross-reference each song once. After creating this generic layout, I proceeded to form comparisons for each song, highlighting similarities and differences in themes, lyrics, storytelling, delivery, and production. Next, I created a Unity script containing a dictionary for each song pair, parallel to its comparison text. Lastly, I created a function that would accept two different songs as arguments, returning the comparison text from the dictionary. The song titles were stored in an Enum for easy selection using UI Dropdown menus. Due to the process of hard coding each comparison, extending this functionality for more songs can be a very time-consuming task. In future iterations, I plan to use OpenAI 's API request to retrieve comparison texts automatically, reducing the manual effort involved in creating comparisons for multiple songs. While I have tested this alternate approach with accurate results, I would also include options for entry of human curated comparisons which would take precedence.

4.7 Lyric Extraction

Additionally, this project includes a lyric extraction feature. This feature gives users the option to query lyrics with keywords to find direct similarities in words and optional synonyms. After supplying text into the input field, users are instantly supplied with buttons displaying the lyric line for each match found. Each button, when clicked, plays the corresponding matched lyric line. To summarize how this feature was implemented, most of this was made possible by using LRC (lyrics synchronization) files. I created a custom parser for Unity to extract the time stamps for each line. Next, I used the current timestamp and following timestamp of each line to calculate the duration of each line. This duration is then used to determine when to play and stop the audio clip of the corresponding lyrics, accommodated by an variable offset time to implement an audio fade. Each song was then parsed into a Scriptable Object containing data for each line. The previously mentioned keyword functionality was then expanded to utilize a collection of parsed songs to query all songs at once. For lyric audio playback, I also included a toggle to isolate the vocals. Overall, I believe that this is a great tool for self-discovery, complimenting the song comparisons matrix.

Figure 6. This image displays the song comparison and lyric extraction features.

4.8 Node Based Interaction

As a final extension, I added a node-based interactions system, allowing users to visually create their own song mixes. This concept was inspired partly by the use visual scripting, which has made much of the data to UI connections in this project possible in the limited amount of time to develop. The main node block allows users to select a song from the 6 select songs used in this project. The interface affords users with options to playback and even access to stems of the selected song, such as vocals, bass, drums, and others. While stemming of the tracks could have been done using API web request, I opted to do this offline to avoid cost while testing. Other nodes were supplied for environmental sounds. This interacts with the practice of soundscape composition artists who repurposed snippets of sounds with small edits. A last node was created to display lyrics of a connected node. Overall, this node system is meant to be used as a playground to explore new ideas and make connections with elements that may not have been connected in everyday settings.

5. CONCLUSIONS AND FUTURE WORK

With the support of innovative tools and the inspiring force of creativity, this project unites the worlds of visual arts and rap music into a unique interactive experience. Using sound spatialization techniques, users are immersed in a controlled auditory setting that allows for multiple voices to coexist, respectful to each other's contribution to the virtual environment. Appreciation is expressed for the pioneers of hip-hop culture who have constructed spaces for rap music and its artists, which has greatly expanded and continues to do so today.

For future iterations of this project, my goal is to extend this template to provide a platform for other movements, cultures, groups, messages, and voices to be elevated and promoted. There is more room to explore the delivery of symbolic references of rap music. In particular, more in-depth analysis of lyrical and metaphorical word play comparisons amongst artists will provide insights of their personal and shared creative techniques. Multi-user capabilities will provide opportunities for joint and shared experiences. Voice and text chat rooms can also invite interesting discourse around the art and music pieces, changing the dynamics of how we typically experience museums. By providing custom editors with ready-made tools and functions, the barrier of entry can be lowered to give visual and music artists additional unique ways to present their artwork.

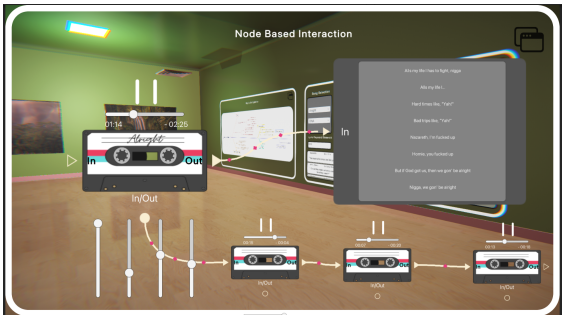

Figure 7. This image displays the node-based interaction remix features.

In conclusion, the integration of rap music in museums, soundwalks, and soundscape compositions presents an exciting opportunity to explore new dimensions of artistic expression and storytelling. By embracing rap music, these cultural spaces can foster inclusivity, amplify marginalized voices, and create dynamic experiences that resonate with diverse audiences.

Acknowledgments

I would like to thank my family for their continued support throughout my academic journey. I would also like to thank the Georgia Institute of Technology, School of Music faculty, and the I.M.P.A.C.T. Lab (Interactive Music, Performance, Audio, and Creative Technologies) for guidance and supporting this project.

6. REFERENCES

[1] B. Yucel, "The Interaction of Painting and Music," Nov. 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877042812032831

[2] "The Hip Hop Museum - The Home of Hip Hop History." [Online]. Available: https://uhhm.org/homepage/

[3] A. d. Paor-Evans, "The Intertextuality and Translations of Fine Art and Class in Hip-Hop Culture," *Multidisciplinary Digital Publishing Institute*, vol. 7, no. 4, pp. 80–80, Nov. 2018. [Online]. Available: https://doi.org/10.3390/arts7040080

[4] J. W. Richardson and K. Scott, "Rap Music and Its Violent Progeny: America's Culture of Violence in Context," *Howard University*, vol. 71, no. 3, pp. 175–175, Jan. 2002. [Online]. Available: https://www.jstor.org/stable/3211235?origin=crossref

[5] C. T. Powell, "Rap Music: An Education with a Beat from the Street," *Howard University*, vol. 60, no. 3, pp. 245–245, Jan. 1991. [Online]. Available: https://doi.org/10.2307/2295480

[6] A. Leach, ""One Day It'll All Make Sense": Hip-Hop and Rap Resources for Music Librarians," Sep.

2008. [Online]. Available: https://doi.org/10.1353/not.0.0039

[7] J. Bailey, "The Cultural Impact of Kanye West," Jan. 2014. [Online]. Available: https://doi.org/10.1057/9781137395825

[8] D. W. Paquette, "Soundwalking and the Bodily Exploration of Places," *Jilin Academy of Agricultural Sciences*, vol. 37, no. 1, pp. 135–145, Apr. 2012.

[9] C. Summers and M. K. Jesse, "Creating immersive and aesthetic auditory spaces in virtual reality," Mar. 2017. [Online]. Available: https://ieeexplore.ieee.org/document/7938144/

[10] H. Westerkamp, "Linking soundscape composition and acoustic ecology," *Cambridge University Press*, vol. 7, no. 1, pp. 51–56, Apr. 2002. [Online]. Available: https://www.cambridge.org/core/journals/organised-sound/article/abs/linking-soundscape-composition-and-acoustic-ecology/FBE211836B584EB92E960B10392E5A5B

[11] PRAKSIS Oslo, "Listening to Rivers - with Leah Barclay and Annea Lockwood," Oct. 2020. [Online]. Available: https://www.youtube.com/watch?v=dGAcUChycp4

[12] "Soundscape Composition." [Online]. Available: https://www.sfu.ca/~truax/scomp.html

[13] musicaincidental1984, "Beneath The Forest Floor -Hildegard Westerkamp," Jan. 2020. [Online]. Available: https://www.youtube.com/watch?v=n-U8iTyqKRA

[14] B. Truax, "Genres and techniques of soundscape composition as developed at Simon Fraser University," *Organised Sound*, vol. 7, no. 1, pp. 5–14, Apr. 2002.

[15] "FlowCanvas | Modern Visual Scripting for Unity." [Online]. Available: https://flowcanvas.paradoxnotion.com/

[16] T. Swiss, "Black Noise: Rap Music and Black Culture in Contemporary America. By Tricia Rose. Hanover & London. Wesleyan University Press, 1994. xvi + 241 pp." *Cambridge University Press*, vol. 14, no. 1, pp. 135–136, Jan. 1995.

[17] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," in *Proceedings of the 40th International Conference on Machine Learning*, ser. ICML'23, vol. 202. JMLR.org, Jul. 2023, pp. 28 492–28 518.

[18] M. Obrist, G. Boyle, M. v. Brakel, and F. Duerinck, "Multisensory Experiences & Spaces," Oct. 2017. [Online]. Available: https://dl.acm.org/doi/10.1145/3132272.3135086