Mapping Sound and Space: A Physically Traversable Audio Corpus Exploring Personal Experiences of COVID-19

RORY HOY, York University, Canada

This paper explores the development process and staging of an architectonic media intervention (AMI) entitled *The Hours Lost*. *The Hours Lost* places visitors within a navigable space populated by a virtual corpus of audio units, mapping their physical movement within the space to the aural playback of these units. The project thematically centres around personal experiences of COVID-19, providing an outlet for public sentiment, recollection, reflection, and projection towards the future following the globally shared cultural disruption of the pandemic's wake. The piece consists of three structures spaced equally within the environment, aimed at a varied temporal frame (past, present, and future). Navigating into one of these structures results in the playback of a complete recorded experience from a "character", a chosen participant sourced before the staging of the work. Visitors may also participate and provide recordings of their own experience within each structure, which are then analysed to become new traversable sonic material for the audio corpus. *The Hours Lost* bridges the virtual and physical within the visual realm through its accompanying AR app, allowing visitors to see the grain-like structures of audio they are triggering, along with a representation of the viral genetic data within the space. Related work aligned with mapping movement to sound, interactive spatial audio, and sonic architectural aesthetics will be presented. The central sonification infrastructure of the work is achieved with IRCAM's CataRT, a concatenative synthesis engine. Various mapping strategies for participant interaction and corpus spatial layouts will be discussed, outlining design decisions for key thematic and participatory realization of the work. Aesthetic considerations of the work will be explored through the lens of microsound composition, sound as aural architecture, collage, and glitch. The therapeutic qualities of the project as a shared communal work will be addressed through current research on expressive therapies, and digital memorialization. Future work will be proposed which theorizes on potential extensions of the AMI, with focus on deepened player interaction, tracking, and mapping strategies.

CCS Concepts: • **Applied computing → Sound and music computing**; *Media arts*.

Additional Key Words and Phrases: concatenative synthesis, aural architecture, expressive therapies, architectonic media interventions

Reference Format:
Rory Hoy. 2021. Mapping Sound and Space: A Physically Traversable Audio Corpus Exploring Personal Experiences of COVID-19. (2021), 8 pages.

1 INTRODUCTION

Interactive media works can allow for the bridging of roles between the participant, author, and spectator, affording those who may have had no stake in the creation of a piece the ability to partake in its unfolding or alter its behavior. This paper will explore the thematic creation process, technical implementation, and resulting form of one such interactive work entitled *The Hours Lost* [10]. *The Hours Lost* is an installation centred around personal experiences of COVID-19, providing an outlet for public sentiment, reflection, and projection towards the future following the globally shared cultural disruption of the pandemic. Movement within the space is mapped to generated audio; fragments of recordings frozen in space, only activated by the passing of visitors. *The Hours Lost* was developed within the context of the graduate research creation class "Vertical Studio Lab" at York University, which called for students of varied research interests to collaborate upon

Author's address: Rory Hoy, rorydavidhoy@gmail.com, York University, 4700 Keele St, Toronto, ON, Canada, M3J1P3.

an "Architectonic Media Intervention" (AMI). AMIs are works which lie at the intersection of architecture and computational art, to question, challenge, and investigate the outcomes of systems and forms upon a site. In its current state, the project is presented as an online demonstration of the intended physical world installation which will follow pandemic restrictions.

A key thematic inspiration for the work is to combat the negative mental toll caused by the pandemic, and to highlight its effect upon the collective population through the presentation of recorded vocal experiences. Panchal et al. [16] report that that the number of adults reporting symptoms of anxiety and/or depressive disorders in the US has risen from 10% to 41.1% from January 2020 to January 2021 - a drastic jump which may have severe impacts upon both the current health landscape and future of our world. Addressing this need for expression of public sentiment, therapeutic output, and communal unity, the project hopes to provide a space for those who are both suffering, or may be unaware of the effects that are going on around them. This need is addressed through the mapping of movement to sound, a semi-guided expression of experiences for participants, and actively engaging in a shared dialogue made possible through a collective space.

Related work will be explored within the realm of relevant sonification techniques, embodied movement sonification, sound's relation to space, and expressive therapies. Thematic goals, aesthetic considerations, and outcomes of the system creation will be presented alongside relevant concepts from the field of sonic arts. The technical implementation of the concatenative synthesis system will be described, along with the key factors for the employed corpus plotting scheme. Finally, future goals will be described for various aspects of the work stemming from its current implementation.

2 RELATED WORK

2.1 Embodied/Movement Sonification

Sonification of movement data is a rich field with a wide array of artistic applications and potential implementations within an interactive environment. A key technological and methodological influence for the project stemmed from the generation and mapping of a "sound space" from Schwarz [22]. The paper discusses CataRT, a concatenative synthesis system allowing for expressive, live, reactive, and flexible generation of audio from an associated established corpus. Concatenative synthesis is a generative audio process in which small segmented fragments of audio are segmented from a provided track. These units are then compiled into a corpus, and plotted on an XY graph. This technique is similar to granular synthesis, in which small pieces of sound are played back by scrubbing through a main audio source. In a typical granular system, this scrubbing is often done as a linear, sequential, or stochastic movement across a buffer, resulting in a range of perceptual results depending on source material, grain size, grain rate, window size, and additional factors which can alter aspects of the sound (such as attack, sustain, delay, and release). Depending on these settings, generated audio can differ drastically from very sparse discrete events to continuous drones, textural shifts, and tones. Xenakis [25] first outlined the musical technique following the theoretical presentation of "quanta" in Gabor [7], with the first real time implementation from Truax in 1986, two years prior to the paper outlining the technique [23]. These systems do not inherently deal with the sonification of movement data, but do allow for flexible mapping of multiple streams/sources as control input to a theoretical limitless audio space (given a corpus which can be changed and other parameters tweaked). The aesthetics of these systems will be further discussed at the end of the following section.

"Embodied sound design" is explored in Monache et al. [5], describing a theoretical and technical mapping process for new sonic output from vocal and gestural input. This scheme entitled "vocal sketching", extends the expression of vocalizations beyond the sounding capabilities of the human

body, and returns potential audio matches based on input. The SEeD (Sound Embodied Design) system created by Monache et al. allows for participants to imitate natural sounds, which become the search criteria for matching within a database/corpus of preaquired audio - effectively mapping the voice as a traversal method through a concatenative synthesis system. While there is a clear departure between physical movement in space and vocalizations, this mapping of action, analysis, and retrieval has artistic parallels to our intention for allowing experiential vocalizations within various spaces of the installation, which are then presented as new material which can be experienced by the traversal of physical (or virtual) space. The development of this sound space leads to questions about the ability to learn a space through sound and the role of mapping physicality to generated audio. Effenberg [6] presents findings on the efficacy of learning when pairing movement and sonification within a motor control efficacy context. The article displays data in regard to matching audio cues to material and kinetic components of an action, and the resulting increase in accuracy of recognition, and recreation of these actions.

2.2 Expressive Therapies

Malchiodi [13] explores "Expressive Therapies", the potential for artistic output and development as a strategy for the reconciling of various emotional, mental, and social issues which people face everyday. The publication demonstrated that providing an expressive output for these issues has been shown to remedy the negative effects of many of the maladies that continue to place strain on communal and personal relationships. Therapies are explored through the process of visual art, music, movement/dance, drama, poetry, and play. Given the participatory nature of the piece and potential for grief due to the stressful content explored in the work, research on digital memorials by Moncur and Kirk [15] and audio memorials by Schmidt [20] became salient in the interaction and thematic design process. Moncur and Kirk describe a framework for the inclusion of viewers, inputs, form, and message within a space of a memorialization as parameters which sculpt the intent and execution of a work (eg. context, desired content, static vs evolving, cultural vs personal, religious vs secular). Schmidt [20] details works which are dedicated as memorials to groups who have suffered acts of political violence, loss, and ethical mistreatment. The composition of these works are framed as not fulfilling the role of religious or liturgical works, but may be performed through a religious or ritual like manner.

3 THEMATIC, AESTHETIC, AND CONCEPTUAL METHODOLOGY

The artistic interaction focus for the development of the audio system within *The Hours Lost* was the generation of a sonic landscape which was traversable through physical movement, with position being the primary point of action and sensing. System design and mapping strategies will be discussed in section 4, while the key aesthetic and thematic goals will be outlined here.

3.1 Sound as Traversable Landscape

Within the space of the piece, there is a conversation established between the physical structures (referred to as loci), the surrounding "empty" space, and visitors who populate the area. This conversation naturally produces hierarchies of control through the affordance or blocking of a participant's ability to move or inhabit certain areas, resulting in natural flows in relation to typical entry zones, and entry and exit points of each locus. The design of the physical forms of each loci will not be outlined here but will be included visually for reference (Fig. 1).

As the piece deals directly with time and the perceptions of those providing their experiences of COVID-19 in relation to the past, present, and future, the thematic and aesthetic frame of understanding these recordings to be akin to the "sands of time" was established. Following this, conceptualizing the distribution of these audio grains across the installation space resulted in a

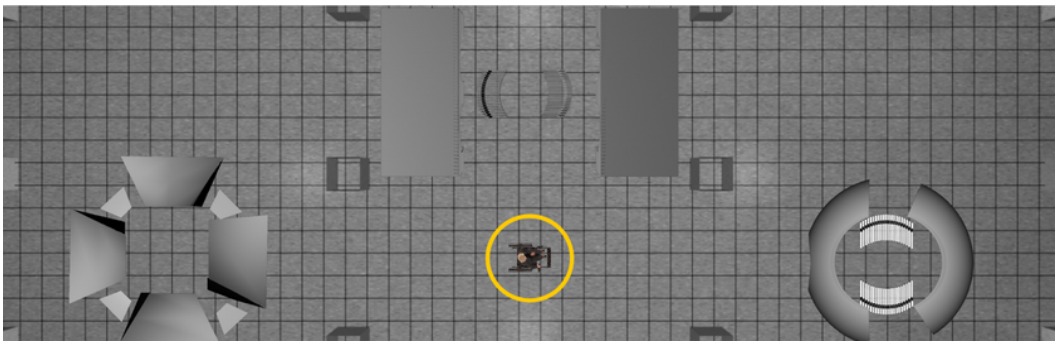

Fig. 1. Plan view of the virtual installation spatial layout

work that links position, time, and sound directly. This symbiosis of physical, temporal, and aural location is addressed through the concept of aural architecture [1]. An aural architecture defines the experience of audio in a space, and the resulting effect of space upon the perception of sound. This may also extend to the consideration of audio within a space as it is constructed. Our understanding of space can be directly tied to our ability to navigate, locate, and orient ourselves in relation to others and structures inhabiting it. The design of a space to incorporate audio adds another orienting perception or dimension which can be used to define the aural structure of a space, and result in a greater understanding of it. The established audio corpus of participant and character recordings is not used or plotted in relation to the loci positions, but the placement of audio units across the installation space results in a cohesive mapping of space and sound. As this plotting does not change in relation to participant input (only added to through additional recordings), it can be used to result in a greater understanding of position in space. Linking this physical location and perceptual phenomenon as a way of learning space is also in line with the previously noted findings of Effenberg [6], in which repetition and understanding of kinetic motor skills or motions are increased through accompanying audio. Participant reactions point to novelty and interest in the ability to revisit a space and re-trigger the same audio; pointing to the musical potential of the audio corpus mapped to physical locations, along with an established memory of location.

3.2 Expression and Catharsis

Providing a space for visitors to reflect, communicate, express themselves, and experience the realities of others, *The Hours Lost* intends to provide a cathartic experience of confronting the challenges and hidden moments through the large population of audio fragments suspended within the space, and complete stories of "characters" which are housed in each of the respective loci. Perception of time has been greatly distorted within the continual lockdowns and monotony of the pandemic [9], which is reflected through the dispersal of audio units in the space. Fragments of stories pertaining to past, present, and future are strewn through, above, and below each other, collapsing time to an equal plane of potential activation. The intent of depicting the displacement and fragility of these stories was to touch upon strands of expressive therapies [13]; drama, music, and movement. The dramaturgical aspects of the work result in a dialogue between space, time, and sound, allowing visitors to express themselves multimodally and experience the expressions of others through a shared space, and temporal flattening. All visitors who provide their own stories further populate the audio corpus with material which can then be experienced in the same space by new visitors at a later time.

Given the potentially solemn nature of the installation's content, design considerations from the Moncur & Kirk [15] framework on digital memorials were employed. As opposed to a static monument or autonomous structure memorializing the stories of the public, *The Hours Lost* is a piece which evolves with the very character of voices from participants, and becomes expressive through the movements of visitors in the space. Visitors are not prompted in any way in regards to traversal of the space and could vary greatly through mobility, mood, and intention. The title of *The Hours Lost* is not direct in its subject matter (similar to pieces discussed within Schmidt [20]), but will be evident upon visiting through the shared cultural experience it portrays.

3.3 Glitch, Collage, and Microsound

Following the description of the audio space as a traversable landscape, this section will outline the aesthetic content of the audio units populating the concatenative synthesis corpus. Ryuichi Sakamoto's "fullmoon" [19] is a key aesthetic inspiration for both the system's content and intention of the piece. Telling the same story through multiple languages, the musical piece moves through varied narrators overlapping and noting the impermanence of time and life. This concept of shared narrative and fleeting moments is at the core of the work, relying on small fragmented units of speech to result in a dense texture of sound.

This compositional technique is explored at various time scales in Roads [18], through outlining techniques such as micromontage, granular synthesis, and varied texture. Roads also outlines the compositional results of heterogeneous vs unique sound materials. As the human voice is relatively homogenous in timbre across all ages and body types within the grand space of potential sound, the material used for the concatenative synthesis system is in turn homogeneous, creating an uncanny valley of understanding from the segmented audio. Clear syllables are discernible, along with a sense of intention and vocal timbre, while all context of meaning surrounding a single unit is lost to the similarity in neighbours of the corpus (which are most likely not placed sequentially from the original source material). This montage of discrete utterances results in a rolling texture following visitor movement, as disjointed or continuous as their input.

The presentation of sound in this disjointed, and textural manor can evoke a sense of glitch aesthetics - works that purposefully use malfunction for artistic output [24]. Classically, glitch-like sounds have been composed using stochastic methods [17] by musicians like John Cage, and Iannis Xenakis, resulting in other worldly, machinic, or uncomfortable effects. This is not the case within *The Hours Lost*, as sounds are fixed and remain stationary in their position - allowing for the theoretical repetition of "play", though this would be difficult given potential inaccuracies in assessing position as well as inaccuracies in human movement.

4 SYSTEM DESIGN OUTCOMES

4.1 Concatenative Synthesis

The primary auditory system of *The Hours Lost* is developed using CataRT [21], a concatenative synthesis engine running in Max 8 [4]. The implemented version of CataRT is created through the use of the Mubu library, developed by IRCAM [12]. To briefly restate the premise of concatenative synthesis from section 2.1, the system conducts audio segmentation upon a source track, resulting in a corpus of small audio units. Upon segmentation, descriptor analysis is conducted, generating a table of mean frequency, energy, periodicity, AC1 (first order autocorrelation coefficient), loudness, centroid, and spread values for each audio unit. This analysis allows for the plotting of grains upon an XY graph which can be traversed to playback the audio content of each unit. Varying the chosen plotting strategy can drastically change the resulting corpus layout and in turn, the character of the traversed audio which will be described in the following subsection. The current working/demo

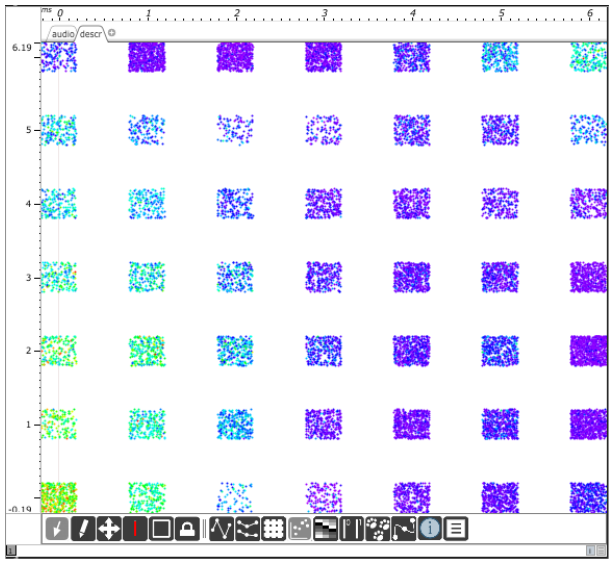

Fig. 2. Defaut SOM layout of CataRT corpus for *The Hours Lost* (colour indicating frequency mean)

audio corpus is populated by roughly 19000 audio units, generated from 44 minutes of sourced character audio. Audio units are segmented every 145ms, maintaining the desired timbral and syllabic content of the voice, while removing contextual and discernible speech. In the ideal physical installation scenario for the work, this resulting audio from movement would be spatialized to speakers closest to the participant using IRCAM's Spat [2]. As others move through the space further from one's immediate area, a dense texture of vocal sounds will be experienced which relate to the approximate position of others being tracked in the space. In this virtual version of the piece, all participants and viewers receive the same audio, resulting in a more closely shared experience but also one which lacks the variability of separate perceptual viewpoints.

Triggering of audio units is done using CataRT's "Fence" method, which results in the playback of the closest audio unit to the input location without repetition. If a participant were to move and the closest unit does not change, there would be no resulting audio. Given the breadth and distribution of the corpus population this is relatively unlikely - but does still occur, resulting in breaks and pauses between discrete units firing. Nearest neighbours are calculated using the mubu.knn object, passing this selection to mubu.concat. This structure is duplicated per participant, allowing each player to have their own voice within the space which does not need to compete for control between visitors.

4.2 Plotting Strategies

A key concern of the corpus generation and plotting scheme was to result in a complete mapping of space to active audio units within the sonic landscape. Typical plotting methods of audio feature XY combinations resulted in interesting and full textural audio content, however there were clear pockets of low audio activity, effectively resulting in dead zones for participants as they moved through the space. This variance in density would be advantageous/desired within a typical musical application of CataRT to allow for clear dynamics in play, however these areas of low activity resulted in a disconnect between action and sounding within our creative context. To combat this, a self-organizing map [14] (depicted in Fig. 2) was employed to distribute the population of audio

units within a defined grid of clusters along the entire corpus space. The artificial neural-net of the self-organizing map produces a lower dimensional (2D) representation of the entire descriptor analysis space listed in the previous section. While this spatial distribution method does result in an even displacement of positions along the entire corpus space, placement into each cluster is not defined by an arbitrary amount or ratio to the complete population, but by the assignment of the SOM output in relation to each input dimensional factor. Because of this, there will still be areas of the space which are perceivably less active, but the main concern of empty pockets has been avoided.

Mapping of input parameters is key within a typical electronic musical system, determining the essence of the instrument through careful consideration of affordances and immutable characteristics [11]. This "bringing of interface to life" results in compelling play dynamics and output. In this non-traditional musical case, the plotting of sonic material/audio units to the physical space acts as our mapping. As the audience is also the performer of the space, participant engagement [8] was considered in relation to varied plotting schemes, physical layouts, and aesthetic goals.

5 FUTURE WORK/CONCLUSION

As the current mapping of movement to sound is linear in relation to the chosen plotting scheme upon the installation space, varied sonification behavior could be achieved by incorporating dynamic processes to the positions of audio units. One possibility is the propagation or disruption of audio unit locations. This could be achieved by simulating a force acting within a suspended physical simulation of the corpus elements. By adding advanced tracking solutions which could take more body parts into consideration, a complex n-dimensional mapping of factors could be developed to separate areas of the body, position, orientation, and expression. Since the current online/demo version of the piece relies upon standard stereo based audio chat services, there is a limitation upon providing each player a unique vantage point within the space (which would be inherent within a physical staging of the work). This may be addressed through a streaming service such as Icecast [26] or performative solutions like JackTrip [3] to send separate audio streams to all players.

The audio system of *The Hours Lost* was developed to explore the potential of space as a traversable sonic medium, foster positive communal and personal communication in the wake of a global cultural disruption, and engage visitors through a unique multi-sensorial installation. Responses from participants point to the novelty and musical qualities of revisiting certain areas in the space, and exploration into the use of vocals as audio units to frame the space as a microcosm of the vast array of experiences of COVID-19 has proved effective both aesthetically and technologically.

REFERENCES

[1] Barry Blesser and Linda-Ruth Salter. 2009. *Spaces Speak, Are You Listening?: Experiencing Aural Architecture.* The MIT Press.

[2] Thibaut Carpentier, Markus Noisternig, and Olivier Warusfel. 2015. Twenty Years of Ircam Spat: Looking Back, Looking Forward. In *41st International Computer Music Conference (ICMC).* Denton, TX, United States, 270 – 277. https://hal.archives-ouvertes.fr/hal-01247594

[3] CCRMA. 2008. JackTrip. https://ccrma.stanford.edu/software/jacktrip/

[4] Cycling '74. 2018. What is Max? https://cycling74.com/products/max

[5] Stefano Delle Monache, Davide Rocchesso, Frédéric Bevilacqua, Guillaume Lemaitre, Stefano Baldan, and Andrea Cera. 2018. Embodied sound design. *International Journal of Human-Computer Studies* 118 (Oct. 2018), 47–59. https://doi.org/10.1016/j.ijhcs.2018.05.007

[6] Alfred Effenberg. 2005. Movement Sonification: Effects on Perception and Action. *Multimedia, IEEE* 12 (May 2005), 53–59. https://doi.org/10.1109/MMUL.2005.31

[7] D. Gabor. 1947. Acoustical Quanta and the Theory of Hearing. *Nature* 159, 4044 (May 1947), 591–594. https://doi.org/10.1038/159591a0 Number: 4044 Publisher: Nature Publishing Group.

[8] Visda Goudarzi and Artemi-Maria Gioti. 2016. Engagement and Interaction in Participatory Sound Art. Hamburg, 6.

[9] E. Alison Holman and Emma L. Grisham. 2020. When time falls apart: The public health implications of distorted time perception in the age of COVID-19. *Psychological Trauma: Theory, Research, Practice, and Policy* 12, S1 (2020), S63. https://doi.org/10.1037/tra0000756 Publisher: US: Educational Publishing Foundation.

[10] Rory Hoy, Filiz Eryilmaz, Raechel Kula, Zian Liu, and Jonathan Silveira. 2021. The Hours Lost.

[11] Andy Hunt, Marcelo M. Wanderley, and Matthew Paradis. 2002. The Importance of Parameter Mapping in Electronic Instrument Design. In *Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02)*. Dublin.

[12] IRCAM. 2008. MuBu. https://forum.ircam.fr/projects/detail/mubu/

[13] Cathy A. Malchiodi (Ed.). 2005. *Expressive Therapies*. Guilford Press, New York.

[14] Jonas Margraf. 2019. *Self-Organizing Maps for Sound Corpus Organization*. Ph.D. Dissertation.

[15] Wendy Moncur and David Kirk. 2014. An emergent framework for digital memorials. In *Proceedings of the 2014 conference on Designing interactive systems (DIS '14)*. Association for Computing Machinery, Vancouver, 965–974. https://doi.org/10.1145/2598510.2598516

[16] Nirmita Panchal, Rabah Kamal, Cynthia Cox, and Rachel Garfield. 2021. The Implications of COVID-19 for Mental Health and Substance Use. https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/

[17] Alexandre Popoff. 2013. John Cage's Number Pieces as Stochastic Processes: a Large-Scale Analysis. *arXiv:1311.5853 [physics]* (Nov. 2013). http://arxiv.org/abs/1311.5853 arXiv: 1311.5853.

[18] Curtis Roads. 2001. Aesthetics of Composing with Microsound. In *Microsound*. MIT Press, Cambridge, 325 – 348.

[19] Ryuichi Sakamoto. 2017. "fullmoon" (from "async"). https://www.youtube.com/watch?v=UyK1VYUY_BI&feature=emb_title

[20] James Schmidt. 2010. Cenotaphs in Sound: Catastrophe, Memory, and Musical Memorials. *European Society for Aesthetics* 2 (2010), 25.

[21] Diemo Schwarz. 2005. Current Research in Concatenative Sound Synthesis. (2005), 4.

[22] Diemo Schwarz. 2012. The Sound Space as Musical Instrument: Playing Corpus-Based Concatenative Synthesis. In *New Interfaces for Musical Expression (NIME)*. Ann Arbour, United States, 250–253. https://hal.archives-ouvertes.fr/hal-01161442

[23] Barry Truax. 1988. Real-Time Granular Synthesis with a Digital Signal Processor. *Computer Music Journal* 12, 2 (1988), 14–26. https://doi.org/10.2307/3679938 Publisher: The MIT Press.

[24] Janne Vanhanen. 2003. Virtual Sound: Examining Glitch and Production. *Contemporary Music Review* 22, 4 (Dec. 2003), 45–52. https://doi.org/10.1080/0749446032000156946

[25] Iannis Xenakis. 1971. *Formalized Music: Thought and Mathematics in Composition*. Pendragon Press, Stuyvesant, NY.

[26] Xiph.org. 2004. Icecast. https://icecast.org/