eLabOrate(D): An Exploration of Human/Machine Collaboration in a Telematic Deep Listening Context

Rory Hoy and Doug Van Nort

DisPerSion Lab - York University
hoy@yorku.ca, vannort@yorku.ca

Abstract. *dispersion.eLabOrate(D)* is a networked performance system which augments and supports Deep Listening workshop experiences through an environment that integrates human and machine collaboration. The sonic materials for this co-performance/creation are seeded by vocal activity of human participants, which continually contribute to an audio corpus of past content used for resynthesis of machine voices. Each participant experiences their own spatial sonic reality within a shared virtual audio space, as relative placement to other collaborative sources provide a unique vantage point via an accompanying virtual acoustics system. Responses from public play sessions are analyzed using a grounded theory approach to report on salient qualitative data resulting from performances with the system.

Keywords: Telematic Performance, Networked Audio, Interactive Agents, Deep Listening

1 Introduction

Network-based communal activity and connection in the area of telematic music grew dramatically throughout the early lockdowns and cancellations caused by the global COVID-19 pandemic [13], as musicians turned to software solutions in order to continue their regular performance sessions at a distance. For example, in this period via the DisPerSion Lab we produced over forty telematic performance events involving more than forty performers. Research and development in our current "post-pandemic" context continues to foster distributed musical practice and telepresence, for ourselves and many others, through various systems capable of very low latency and high quality audio. One distinct performative practice that was impeded by the lack of in person events, both in our own local lab context and more broadly, was Deep Listening – described as "a practice that is intended to heighten and expand consciousness of sound in as many dimensions of awareness and attentional dynamics as humanly possible" [18], by its creator Pauline Oliveros. Public engagement with the typically in-person and group-based Deep Listening workshop events were therefore put on hold until restrictions had lightened.

Continuing this trajectory of research-creation work conducted on telematic performance, Deep Listening practice and human/machine co-creation, the system outlined in this paper was developed to explore the results of augmenting a well-known Sonic Meditation-style [20] text piece by Oliveros from the Deep Listening literature called The Tuning Meditation. In particular, this augmentation introduces non-human agents directly performing alongside human participants. Entitled *dispersion.eLabOrate(D)*, this project is a conceptual reconstruction and extension of a previous system by the authors, entitled *dispersion.eLabOrate* [10] - with the updated name reflecting that the system, and the practice that it fosters, is now *(D)istributed*. Vocalizations made by human performers during this piece are captured in an audio corpus which is then used both as raw material for the synthesis of new tones by the machine agents, as well as functioning as their running memory of past sonic events. The population of agents is variable and each acts autonomously according to the score for The Tuning Meditation (which will be outlined in section 2). While physically dispersed, players and machine agents are placed within a singular virtual acoustics space to be heard within the same environmental conditions.

One challenge for this project, something faced by most telematic-based performances, is the collapse of spatial qualities to a (typically) stereo mix of all performers. This flattens any variance in positioning of local sound sources one would find within an in-person event, which can provide contextual information or sonic material to react to. Systems that can support spatial representations of source placement typically orient sound to a particular "sweet-spot" in the centre of the virtual space, which all sources are placed relative to. For *dispersion.eLabOrate(D)*, we develop and present a spatial audio setup which allows for unique sonic perspectives tied to relative placement within the virtual acoustics environment.

Following the completion of the system, Deep Listening workshop sessions were held to gather qualitative feedback on various elements of the experience. These responses are presented and analyzed with a grounded theory approach in section 5. Key categories of responses are discussed, which were found to focus on immersion and communal space, diversity of machine voices, and strategies for human/machine collaboration.

2 Related Work

Our previous work on augmenting Deep Listening practices emerged over the course of a 12-week DisPerSion Lab seminar that posed the question: "Can we imagine ways that interactive systems might synergize, entangle with, and augment – but not distract from – Deep Listening practice?" This resulted in our performance system *Dispersion.eLabOrate*, created for collective listening and sounding in a shared physical space. This system can detect and react to player vocalizations as well as ambient sound within an environment. Like this newer project, *eLabOrate* was designed to engage and augment group performances of The Tuning Meditation (TM), placing the output of the system as a machine agent which engages with the vocal and collaborative dynamics of the human participants. The TM asks participants to focus on their breath – inhale deeply, exhaling on a tone of their choice for one full breath. On the following exhalation, match a tone currently being sung by another player. Then on the next vocalization, sing a new tone that hasn't been sung yet. This cycle continues until a natural end point is reached where each player has stopped.

Building upon this past work, we once again begin from the position that the machine voice/participation is not an element which should conceptually or perceptually dominate the piece, but rather should work in tandem alongside human performers to facilitate broadened sonic potentials. We investigated the past system through the lens of "Sonic Ecosystems", foregrounding resonance, feedback, and autonomous behaviour inside/outside the direct influence of human action. Building upon related works in the field, Sonic Ecosystems are framed here as performative contexts – ones which establish environments that in turn adapt to agents, which define their own self-regulating populations and their own ambience. This ambience may be naturally-occurring within the acoustic space facilitating the system, or could be generated as a result of the sonic ecosystem's behaviour. These systems rely on self-monitoring techniques both virtually and physically, often including microphones or other sensing devices within the space in order to enact and react to these recurrent activity loops [6]. Musick (2016) [16] provides a thorough look at the theory and practice of the field within their Sonic Space Project, and includes assessment strategies of sonic ecosystems within their 2014 paper [15].

In addition to *eLabOrate*, another recent DisPerSion Lab project by Maraj and Van Nort [14] also focused on developing an agent-based system for interactive performance, building upon rules found in a Sonic Meditation-style piece. In this case the text piece Interdepence was used as a starting principle for structuring agent interaction, and the focus was on gestural performance with an interactive system rather than collective vocalization in a workshop setting. Both projects emerge from and sit at the intersection of two larger DisPerSion Lab projects that engage this broader area, entitled Deeply Listening Machines and Deep Listening Entanglements [7], both of which build upon past work on intersecting machine improvisation and Deep Listening principles [24].

3 System

eLabOrate(D) allows for more complex, refined, and flexible agent behaviour to that of its predecessor *eLabOrate*. Where *eLabOrate* created a pervasive and mirroring-like behaviour for all vocalizations and ambient sound, *eLabOrate(D)* more closely follows the cyclical behaviors of new tones, matching, and most importantly active listening as is requested within the context of The Tuning Meditation. A key concept discussed within the Deep Listening community of practice is the distinction between directed and focused active listening, as opposed to the passive physiological process of hearing [19]. The behaviour for each machine agent in the system enacts this active listening as opposed to the more passive and reflexive hearing and sounding which occurred in our past work. Situating this performance system as a telematic piece, we also investigate the viability of a remote virtual shared space as a facilitator for the characteristics and behaviour of sonic ecosystems.

In practice, the *eLabOrate(D)* system is comprised of modules to capture human signals and generate machine voices created in Max/MSP, which are detailed in the following subsections. These are instantiated at the beginning of a session, with an individual module for each human/machine participant being scripted once the total number of participants is selected. These modules allow captured and generated sound to be sent to an accompanying patch to be spatialized appropriately for each human participant. All participants (both human and machine) are placed in a circle within this virtual acoustics space, which is relevant for machine voice behaviour and will be outlined in subsection 3.2.

3.1 Audio Corpus - Sonic Memory

The audio corpus is implemented through the Mubu package [21] in Max, and is used within the system as a running memory of all vocalized tones. Input signals from human players are segmented into audio buffers based on detected onsets, with recording taking place until the signal falls below an established threshold. The system is limited to the last n (default 50) segmented tones in order to avoid memory & processing issues, but could be extended depending on the hardware capabilities of the host computer. This places the corpus as a short term memory of sonic events, as the oldest events are erased when a new buffer is saved to the corpus beyond the limit. Buffer input is held for a short time after vocalization ends, to allow machine voices the opportunity to match without the target voice being explicitly active. This behaviour is similar to human participants matching another tone briefly after a given vocalization has stopped, which happens often in practice. The buffer content is analyzed and segmented through Mubu and is made accessible within the corpus as separated grains.

3.2 Machine Module

Each machine module consists of separate logic sub-modules inside. Controlling the movement between matching and new tone behaviors is the breath control module, which mimics a range of human time scale breath cycles (Fig. 1). This approximation of human breath allows the machine agents a voicing and breathing alternation, so as to both avoid continuous output and to better align with the time scales present within a typical performance of the Tuning Meditation.

New tones are made up of grains from the collective audio corpus populated by human participant voices, scrubbed through and resynthesized using a concatenative synthesis [22] method within Mubu. A target frequency area is scrubbed to playback these grains as a continuous voice, with the resulting tone constituting a new voice comprised of grains from various participant sounds.

Machine voices are assigned one of 6 possible states upon instantiation, which define either their matching (Seeker or Buddy) or their spatial biasing (Close, Far, Unbiased) behaviors. Matched tones result from copying a desired target's current voice buffer content into the acting machine voice's buffer. This buffer is then played back with a granular synthesis method to mimic the held tone by another player (both human and machine). The *Seeker* state implements a spatial encoding neural net using

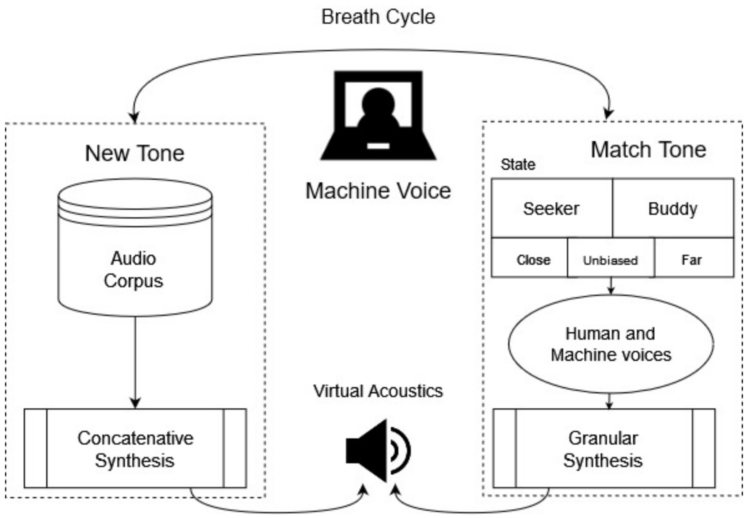

Fig. 1. Machine voice module depicting breath cycle alternating between both vocalization states and synthesis engines

ml.spatial [5] to keep track of recently-matched participants. A recent match of a participant would mean that on the next match behaviour, they become the least likely to be chosen. The *Buddy* state chooses a small subset of the total population to use as potential matching targets. A spatial bias of near/far/unbiased for both *Seeker* and *Buddy* skews the likelihood to match a specific subgroup based on angular proximity within the established virtual circle. Participants within a ±90° degree arc from a given voice or less are deemed "close", while beyond this places a voice as "far". Through these added logic modules, the system possesses possible "listening behaviours" that are above and beyond the basic instructions of the TM piece, thereby providing differing kinds of performative identities for each machine voice.

3.3 Human Module

Human participants are connected to the system and to one another by utilizing JackTrip [1], a software tool for low latency and high quality audio. Depending on the microphone setup of each participant, acoustic characteristics of each player's environment may be sent along with their vocal input allowing for a perceived commingling of ambient sounds within the virtual acoustics space. These various local sonic realities are relayed along with one's intended vocal utterance of a new tone or matched tone, and may therefore become sonic material that a machine voice may pull from to generate its own new tone. In this way, the expanded sense of shared acoustic environments remains present as a point of focus during these sessions - in the spirit of Deep Listening practice that emphases attention to one's sonic environment. These vocalizations can be conceived of as "the performer-instrument articulation", theorized by Waters as

> "[...]result[ing] not only from the physiology of the player, but also from the complex feedback into that player's body of vibrating materials, air, room, and the physiological adaptations and adjustments in that body and its 'software' which themselves feed back into the vibrating complex of instrument and room." [25]

This feedback is additionally facilitated by the networked nature of the system, and in the individual behaviours of each machine voices. Resynthesis of vocal material of participants continually shapes the timbral range of the machine voice, and in turn has influence over the potential tones a human participant may match through the score. This also comes about through the matching behaviour of the machine voices, as they are able to copy another vocalizing machine.

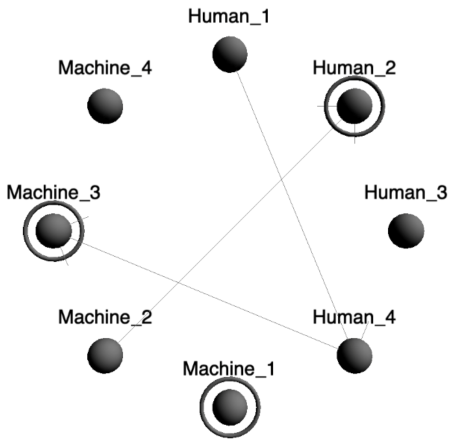

Fig. 2. Layout of human and machine voices (top-down) within the accompanying virtual acoustics space. Lines denote matching behaviour and rings denote new tones being vocalized.

3.4 Multi-Spat - Virtual Acoustics

All output is processed and positioned spatially within an accompanying multi-listener spatialization patch developed for this system. Audio from each human and machine participant is placed at a corresponding source location distributed evenly in a circular pattern around the virtual space. Relative spatial listening mixes are achieved through separate instances of IRCAM's Spat 5 [2], with one virtual space model per human player. In practice, this allows a hypothetical listener 2 to be placed to the left of listener 1, and be heard from that direction by listener 1. The same is true for listener 2; their relative position to 1 would perceptually place listener 1 to their right. Each instance of Spat is run using a binaural panning mode for a stereo output, as headphone based monitoring is encouraged for the performance. Two virtual audio drivers [12] [9] are employed to allow for routing to and from JackTrip for each participant and allows a separate and accurate binaural mix to be passed individually back to players in relation to their virtual orientation. An accompanying visual layout of the spatial positions and behaviour of each participant can be displayed to depict current activity of sources (Fig. 2). New tones are depicted as rings around a given source, and matching behaviors point to the target source a voice is matching.

4 Study/Play Sessions

Telematic Deep Listening workshop sessions were held with various sized groups of players in order to explore multiple factors of engaging with machine agents in this setting. The sessions were facilitated by the second author, a certified Deep Listening instructor, and placed focus on performance of the Tuning Meditation in the context of also drawing attention to the shared sonic environment, one's local environment, one's body and to inner listening - all common elements of a Deep Listening workshop session.

Participants were invited through calls sent out to online email lists and social media groups focused on computer music, Deep Listening, sound art/studies and listening more broadly. Based on scheduling alignments, we arrived at 8 total participants who connected from disparate locations in North America and Europe. Multiple sessions were held, which included an equal number of human and machine players at a given time (eg. 4 human players, 4 machine players in one session). Sessions were an hour in length and included two different performances of the TM, each with a different active state for the machine voices. States were decided randomly (without duplicates) in order to have at least one response to each of the varied matching behaviors across all of the sessions. To recap, these states include:

- Far Seeker, Close Seeker, Non-biased Seeker
- Far Buddy, Close Buddy, Non-biased Buddy

These states introduce a bias towards spatial positioning of participants (far, close, unbiased), and a matching behaviour (Seeker or Buddy) which alters how the machine voices attempt to match another vocalizing participant.

Participants were not primed on several factors of the experience, as we were interested in gathering undirected qualitative data for analysis using a grounded theory approach. Grounded theory is a qualitative methodology aimed at uncovering key information from responses and allowing for central themes to emerge via multiple stages of coding - extracting relevant data and ultimately "Crystallizing the significance of the points" [3]. Our grounded theory-based approach for this study consisted of separate open coding steps each done individually by both authors. After this initial coding pass, cross-checking of codes occurred followed by focused coding - creating larger categories of responses that were synthesized from the resulting codes and will be presented in the following section.

After engaging in two runs of the meditation, players were invited to complete an online form in the (approximately) fifteen minutes remaining within their given session. The questions were designed to allow for open reporting on the experience with the goal of allowing key areas of personal interest and thought processes to come to the forefront. That said, once the session was completed it was clear to participants that there were both human and machine voices present, and so we explicitly asked about the experience relative to these distinct entities. The questions provided were as follows:

1. What are your general thoughts about the session?
2. How would you characterize the various voices (both human and machine)?
3. Could you compare and contrast your experience of the two (human/machine)?
4. What was your strategy for following the TM piece?
5. Could you characterize your relationship to space and describe if (and how) it might have influenced your experience?

To clarify demographic and background information, we also asked participants if they had any previous experience with Deep Listening as a practice, and if they had ever performed The Tuning Meditation.

5 Analysis

Three key categories of focus emerged through our process of coding participant responses and subsequent analysis, which we will discuss in the following sub-sections.

5.1 Experience of Immersion in a Shared Communal Space

A recurring theme that was prevalent throughout participant responses was a sense of immersion, with this being tied to a characterization of the session as a shared communal space. “I lost all sense of my local space. With eyes closed I was entirely in a shared space with everyone. I forgot we were not physically together”, noted one participant with an extensive background in Deep Listening practices. As one might expect, specific mention of the term “sonic ecosystems” was not present in responses, however those characteristics we previously identified as belonging to sonic ecosystems were indeed reported, with one player explicitly noting “I noticed a bit of a back and forth between being influenced by the machines and the other humans in the session”. In response to a subsequent question, they evoked metaphors of physical ecosystems:

> “I was visualizing the sound itself a lot more. I felt like I was contributing to a moving stream. I didn’t know how loud I was, and so it felt like I was occasionally throwing a bucket of dyed water into the stream as the water flowed by, changing it in ways I wasn’t aware of.”

This ambiguity of outcome also gestures to the lack of direct control over the system. While each performer is an active participant in seeding the amalgamated voices generated by the audio corpus, there is a blurring of causal human action to machine reaction. This is congruent with the concept of a “floating phenomena/floating piece of art” from Weibel & Dinkla, described by Dixon as “[...] no longer the expression of a single individual. Neither is it the expression of a collective, but it is the state of a ‘connective’ - a web of influences that are continually reorganized by all participants.” [8] This is further emphasized by another participant, in commenting on the influence that the space had on their interaction with others:

> “The relationship with space was expansive, I was traveling across the space to meet the tones. It highly influenced my experience particularly in the second tuning. I was able to go with more ease and pleasure, as the space expanded for me, on the possible tones and voices to tune in.”

These performer statements, representative of the broader viewpoints expressed, depict experience within the session as taking part within a shared space which affected their perception of the inter-relational action present within the Tuning Meditation.

5.2 Diversity of Machine Voices Expanding Timbral Content of Meditation

The sense of a diversity of voices from machine agents was reported, and this was characterized as allowing for extended timbral content beyond the human. One participant stated: “[...] it was a great experience to listen to both human and machine voices and respond to them in real time, reflecting my own impression on them. Analyzing various notes of multiple voices was not an easy task but I enjoyed finding atonal harmony in inharmonious sounds.” Similarly, conception of collective space was also addressed via perception of the machine voice character, as one participant articulated: “The expansion of tones was really helping me to expand my sense of space and time, and my connection with others, and with my body tuning in.” This reinforces sentiment from the previous subsection (5.1) while here being expressed in relation to the “expansion of tones” offered by the machine voices.

One participant characterized the voices taking part within the piece as “diverse, some calmer than others. Most of them steady, but some evolving and agile”, which highlights the varied approaches that both human and machine voices took in either matching or new tone vocalizations. Another participant expressed that the character of the machine voices was “[...] radical, refined, with a different atmosphere, pleasant too in a different way.” These statements cause us to question if this evolutionary/radical character was a product of the cyclical matching behaviours that are capable within the system, depending on the nature of the machine voices’ behavioural states. As one voice matches another, a chain-like effect can occur between both human and machine players that either match that same voice, or a voice that is already matching another (visible in Fig. 2). Once established, this type of chain may build upon and subtly (or not so subtlety) alter the timbres at play and seed new material into the audio corpus which defines the machine voices. If and when this is established, such diversity and “refined” nature of the machine voices may also be a direct result of the concatenative synthesis technique used to derive the new machine tones via the “raw material” of the captured player’s voices.

In this complex human/machine network we can only speculate on the specific causalities - though we do note the above as affordances of the *eLabOrate(D)* system that we know to be at play in establishing a sense of evolution and refinement over time. What we can say with more certainty however, is that the machine voices were characterized as diverse, refined and evolving, both in terms of their timbral character and in their behaviours of interaction. This was articulated as a central influence in moving the dynamics of human attention forward in time.

5.3 Reflexive Engagement with Machine Voices and Influence Upon Performance Strategies

An openness towards the involvement of the machine agents was clear in the previously-mentioned responses. For some participants, the overall experience was that all audio was collapsed into a cohesive sounding body: "At a few points, I couldn't tell which were human and which were not."

When they are clearly recognized, the incorporation of accompanying machine agents introduces such Deep Listening sessions to sonic material and gestures which would not occur within a human-only performance. It was reported by some participants that this coaxed out playful transgressions upon the score itself, with one player noting: "I tried to make sound[s] that were machine-like myself. I tried to 'chop' my voice, by tapping my cheek or throat, as I was influenced by the other sounds that were made." Such transgressions certainly can (and do) come about in general during pieces such as the TM by participants who would like to push the limits of what constitutes a held tone, a pitched sound, etc. – and this is an important aspect of the social dynamics of this and similar pieces. Through the interjected behaviours of the machine voices, new timbral and rhythmic components are often introduced into the palette of materials which participants are engaging with. This added dimension of uncertainty and dynamism from this hybrid context is captured by Waters, who states "One of the benefits of hybrid (physical/virtual) systems is their very impurity: their propensity to suggest or afford rich unforeseen behaviors which engage the player (and the listener) at a variety of levels: sonic, tactile, and dynamic." [25].

For some respondents, the incorporation of machine voices changed the conception of their own voice in relation to others. One participant stated, "The machine tones bring a strength from me, fearless voicing. The human voices invite me to listen more, and to engage with care for them, trying to explore the soft voices I haven't listen[ed] to yet." In contrasting human and machine contributions (Q3), another respondent expressed, "The human is easier to follow accurately, the machine leaves more room for interpretation of the note and timbre (which I enjoyed!)", further relating these characteristics to one's own strategy for realization of the score's instructions.

6 Conclusion & Future Work

We have presented our system *dispersion.eLabOrate(D)*, a set of telematic autonomous participants who engage the Tuning Meditation in the context of a Deep Listening workshop setting. An overview of our previous research and a system overview were presented, providing context and structure for relating the design of the system to a set of broader reflections, which were informed by a set of qualitative data that emerged from a series of workshop sessions with the system. Through these sessions, we investigated the potentials for augmenting group Deep Listening practices in a telematic setting via machine participation, and presented an accompanying virtual acoustics system allowing for unique sonic vantage points into the collective virtual performance space.

Responses to play sessions were approached and processed through a grounded theory methodology to parse out key "codes" and broader themes, resulting in the three salient response categories outlined in section 5. These revealed a sense of immersion that was tied to “space”, understood as dual conception that was both social and sonic, a sense of evolving diversity that was carried forward by awareness of machine actors, and a set of strategies that articulated human responses towards positioning themselves within this sonic-communal engagement.

In future work, we will look to further explore and assess these perceived dimensions in this Deep Listening performative context, iterating our design (both computational and workshop structure) in light of what we’ve learned this far. This includes an examination of the concepts of immersion intensity [4] & presence [23]). From a design perspective, future considerations include new and varied implementations of virtual acoustic parameters, and expanding timbral possibilities concerning voices of the machine agents. Both realistic representations of in-person performance spaces and extended potentials for virtual space (physically impossible listening orientations, source positions, room qualities, etc.) offer new potentials for facilitating Deep Listening practices. This virtualized potential for Deep Listening practice is a key component of the *eLabOrate(D)* project in particular.

More broadly, this work contributes to the larger Deeply Listening Machines and Deep Listening Entanglements lab projects. These sister projects seek to transform and augment the kinds of listening and sounding practices found within the Deep Listening literature, such as that expressed by the Tuning Meditation (or Interdependence, in the case of Intergestura). These existing text scores act as starting points - seed ideas - for an evolving set of structured approaches to collective listening and sounding, both in public workshop and improvised performance settings. Each system such as *eLabOrate(D)* is part of this ecosystem of human/machine engagement in a Deep Listening context. Thus future work for this system is focused on diversity of approaches, such as exploring different methods for machine voice synthesis based on sound analysis and machine learning from human vocal inputs, and on modularity such that these particular machine voices might evolve new listening/sounding rules, and interact with other agents that emerge from the larger project.

References

1. Cáceres, J., Chafe, C.: JackTrip: Under the Hood of an Engine for Network Audio. Journal of New Music Research 39 (3): 183—187. https://doi.org/10.1080/09298215.2010.481361. (2010)
2. Carpentier, T.: A New Implementation of Spat in Max. In: 15th Sound and Music Computing Conference (SMC2018), 184–191. https://hal.science/hal-02094499. (2018)
3. Charmaz, K.: Constructing Grounded Theory. Thousand Oaks, Calif: Sage Publications. (2006)
4. Colsman, A., Aspã., L.: Development of a Questionnaire to Investigate Immersion of Virtual Acoustic Environments. In: DAGA, Aachen, Germany. (2016)
5. Day Smith, B.: Ml.* - Machine Learning Toolkit in Max. https://www.benjamindaysmith.com/ml-machine-learning-toolkit-in-max. (2017)
6. Di Scipio, A.: Sound Is the Interface: From Interactive to Ecosystemic Signal Processing. Organised Sound 8 (December): 269—277. https://doi.org/10.1017/S1355771803000244. (2003)

7. DisPerSion Lab.: DisPerSion Lab Research Areas. https://dispersionlab.org/research/. (2023)
8. Dixon, S.: Digital Performance : A History of New Media in Theater, Dance, Performance Art, and Installation. Cambridge, Massachusetts ; London, England : The MIT Press. http://archive.org/details/digitalperforman0000dixo. 561. (2007)
9. ExistentialAudio.: BlackHole. C. https://github.com/ExistentialAudio/BlackHole. (2019)
10. Hoy, R., Van Nort, D.: Augmentation of Sonic Meditation Practices: Resonance, Feedback and Interaction through an Ecosystemic Approach. In: Kronland-Martinet R., Ystad S., Aramaki M. (eds) Perception, Representations, Image, Sound, Music. CMMR 2019. LNCS, vol 12631. pp. 591–599, Springer, Cham. (2021)
11. Hoy, R., Van Nort, D.: A Technological and Methodological Ecosystem for Dynamic Virtual Acoustics in Telematic Performance Contexts. In Audio Mostly 2021, virtual/Trento Italy: 169-–74. https://doi.org/10.1145/3478384.3478425. (2021)
12. Ingalls, M.: SoundFlower. Objective-C. https://github.com/mattingalls/Soundflower. (2014)
13. Keller, D., Costalonga, L., Messina, M.: Editorial: Ubiquitous Music Making in COVID-19 Times. In: Proceedings of the 10th Workshop on Ubiquitous Music (UbiMus 2020) https://halshs.archives-ouvertes.fr/halshs-03035034. (2020)
14. Maraj, K., Van Nort, D.: Intergestura: A Gestural Agent Based on Sonic Meditation Practices, In: 13th International Conference on Computational Creativity, Bozen-Bolzano, Italy. (2022)
15. Musick, M.: Examining the Analysis of Dynamical Sonic Ecosystems: In Light of a Criterion for Evaluating Theories. In: ICMC 2014, Athens, Greece: 154–161. (2014)
16. Musick, M.: Practice-Led Research / Research-Led Practice Identifying the Theory and Technique of Sonic Space Ecosystems. Ph.D., New York University. https://www.proquest.com/docview/1871306234/abstract/9000417E723B4A90PQ/1. (2016)
17. Neidhardt, A., Schneiderwind, C., and Klein, F.: Perceptual Matching of Room Acoustics for Auditory Augmented Reality in Small Rooms - Literature Review and Theoretical Framework. Trends in Hearing 26 (January): 233121652210929. https://doi.org/10.1177/23312165221092919. (2022)
18. Oliveros, P.: Deep Listening: A Composer's Sound Practice. iUniverse, Lincoln. (2005)
19. Oliveros, P.: Quantum listening: From Practice to Theory (to practice practice). Culture and Humanity in the New Millennium: The Future of Human Values, 27–41. (2002)
20. Oliveros, P.: Sonic Meditations. Baltimore, MD: Smith Publications. (1974)
21. Schnell, N., Röbel, A., Schwarz, D., Peeters, G., Borghesi, R.: MuBu & Friends - Assembling Tools for Content Based Real-Time Interactive Audio Processing in Max/MSP. In: ICMC 2009, Montreal, Canada: 423–426 (2009)
22. Schwarz, D.: Corpus-based concatenative synthesis. IEEE signal processing magazine, 24(2), 92–104. (2007)
23. Støckert, R., Bergsland, A., Xambó, A.: The Notion of Presence in a Telematic Cross-Disciplinary Program for Music, Communication and Technology. In Music Technology in Education, 77—101. Cappelen Damm Akademisk/NOASP. https://doi.org/10.23865/noasp.108.ch3. (2020)
24. Van Nort, D., Pauline, O., Braasch, J.: Electro/Acoustic Improvisation and Deeply Listening Machines. Journal of New Music Research 42 (4): 303–324. https://doi.org/10.1080/09298215.2013.860465. (2013)
25. Waters, S.: Touching at a Distance: Resistance, Tactility, Proxemics and the Development of a Hybrid Virtual/Physical Performance System. Contemporary Music Review 32 (2–03): 119—134. https://doi.org/10.1080/07494467.2013.775818. (2013)