Embodied eTube Gestures and Agency

Tommy Davis
tommy.davis@mail.mcgill.ca
IDMIL/CIRMMT/McGill University
Montreal, Quebec, Canada

Kasey Pocius
kasey.pocius@mail.mcgill.ca
IDMIL/CIRMMT/McGill University
Montreal, Quebec, Canada

Vincent Cusson
cusson.v@gmail.com
IDMIL/CIRMMT/UQAM
Montreal, Quebec, Canada

Maxwell Gentili-Morin
maxwell.gentili-morin@mail.mcgill.ca
IDMIL/CIRMMT/McGill University
Montreal, Quebec, Canada

Philippe Pasquier
pasquier@sfu.ca
Metacreation Lab/Simon Fraser University
Vancouver, British Columbia, Canada

ABSTRACT

Following three years of working on the eTube with a collaborative interdisciplinary team, we describe the embodied and spatialized performance gestures developed in tandem with a microphone setup for interacting with musical improvising agents. Eric Lewis' discussion of the intentional stance and make-believe are outlined as a way to conceptualize our working process and engagement with musical agents in improvisation. Within this context, we consider the various agencies at play, and how the musical agents challenge notions of the social and embodiment in improvisation. We will then describe certain artistic approaches and results that are afforded by merging the philosophical with practice. A collaboration with other artists will illustrate new eTube performance gestures. Finally, we outline spatialization models designed for the musical agents and how these were developed in the context of the eTube performance practice.

CCS CONCEPTS

• **Applied computing** → **Sound and music computing**; • **Human-centered computing** → **Gestural input**; *Collaborative content creation.*

KEYWORDS

Embodied performance practice, gesture, improvisation, musical agents, spatialization, eTube

ACM Reference Format:
Tommy Davis, Kasey Pocius, Vincent Cusson, Maxwell Gentili-Morin, and Philippe Pasquier. 2024. Embodied eTube Gestures and Agency. In *9th International Conference on Movement and Computing (MOCO '24), May 30–June 02, 2024, Utrecht, Netherlands.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3658852.3659084

1 INTRODUCTION

We have designed an augmented wind instrument called the eTube to perform with improvising software which uses a microphone to interface between the performer and machine [8]. The eTube is made from a flexible polyvinyl chloride (PVC) tube, which is lightweight and has a directional sound. These characteristics have proven artistically fruitful for musicians to explore movement and spatialized performance gestures in relation to the onstage microphone. We will discuss the development of a gestural movement and sound practice with this simple wind instrument through interaction with improvising software in musical performance.

Musical metacreation (MuMe) is defined by Philippe Pasquier as a "subfield of computational creativity that focuses on endowing machines with the ability to achieve creative musical tasks, such as composition, interpretation, improvisation, accompaniment, mixing, etc." [28, p. 2:4]. Unlike artificial intelligence (AI) research targetted at problem-solving, there are no "optimal solutions" in MuMe, like in art and improvisation [28, p. 2:2]. Improvising software systems are often described by the term "musical agent" (MA), defined by Kıvanç Tatar and Philippe Pasquier as "artificial agents that tackle musical creative tasks, in part or as a whole, and use the methods of [multi-agent systems] and Artificial Intelligence to automatise these tasks" [34, p. 56]. Agents are autonomous and may decide whether or not to act upon a request from another agent in real-time [40, p. 35]. This means that once started, MAs will operate autonomously by generating context-dependent musical output in real-time based on some kind of stimulus. In contrast, devices like MP3 players are not MAs since they are not proactive, when music is played back on an MP3 player the content is simply reproduced.

George E. Lewis defines improvisation as a "social location inhabited by a considerable number of present-day musicians, coming from diverse cultural backgrounds and musical practices, who have chosen to make improvisation a central part of their musical discourse" [21, p. 234]. In addition to the social aspect, improvised performance is a highly embodied practice [32, p. 1]. This may include communication through physical gestures between improvisers, how one's body is temporally situated in an improvisation, how one experiences sound, the ways one engages with a musical instrument's materiality, and the embodied knowledge of performing that instrument, for example. Working with MAs challenges the focus that the social and embodied occupy in improvisation since one must then ask how the machines enter into this social location as disembodied entities. Indeed, this forces one to think more deeply about what the social is, and also what is important about bodies in improvisation.

MOCO '24, May 30–June 02, 2024, Utrecht, Netherlands

ACM ISBN 979-8-4007-0994-4/24/05
https://doi.org/10.1145/3658852.3659084

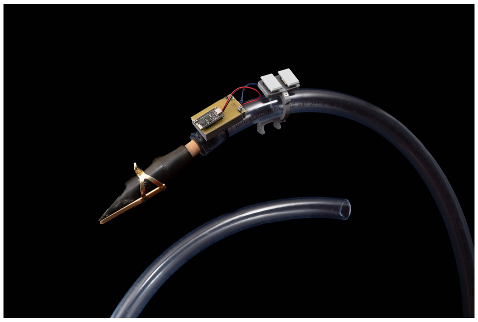


Figure 1: The eTube and two-button controller

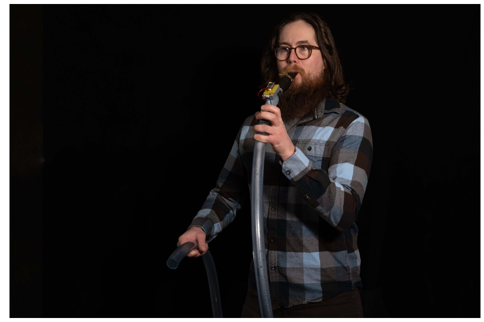


Figure 2: Davis playing a baritone saxophone eTube

We use a research-creation methodology which integrates research and creative expression, blurring the traditional boundaries between the two, while generating knowledge through the act of creation [27]. As we develop this project and investigate relevant research, we reflect on how we conceive of, contextualize, write about, and present the MAs to the public. The main questions we have encountered are: how do we define agency concerning the MAs we use; how does our definition of agency engage with our approach to improvisation; and what creative solutions are afforded by these relationships? We will discuss these considerations in the larger context of an embodied gestural performance practice with the eTube which interfaces with MAs in improvised performance.

1.1 The eTube Project

The eTube is a PVC tube augmented with a two-button controller and capped with a woodwind mouthpiece (see Figure 1). It was first developed by Vincent Cusson and Tommy Davis in 2021 to be performed with the MAs in the *eTu{d,b}e* framework [8]. Davis' performance gestures, improvisation, and movement practice with PVC tubes were first developed during site-specific and improvised dance performances. Here Davis would move and play in the space with dancers, interact with the environment's acoustics, and move the tube for theatrical and acoustic effect. Now with the eTube, these gestures have been adapted into a localized performance practice around a stationary microphone.

The *eTu{d,b}e* improvisation framework uses existing MAs created by other developers in an improvised performance with the eTube [8]. These MAs include the Creative Dynamics of Improvised Interaction (DYCI2) developed by Jérôme Nika and colleagues, Spire Muse developed by Notto Thelle and colleagues, and Construction III created by Sergio Kafejian [16, 25, 26, 36]. Our work with the eTube and these MAs are described in previous publications [7, 8, 29, 30]. Kasey Pocius has developed interactive spatialization models to spatialize the MA's sonic output, which will be described in Section 3.2 [29, 30]. A brief description of DYCI2, and specifically its threshold setting, is necessary to frame our discussion. DYCI2 uses multiple agents which each have their own audio recordings that feed their output. This output is informed by an audio descriptor analysis of a live input. Each agent also has a threshold setting which adjusts the minimum live sound input that is necessary to launch a sonic response from the MA. By adjusting each agent's threshold levels, we may influence the intensity and density of the MA's output. We will focus on how DYCI2's threshold settings have influenced Davis' performance practice and interaction with the microphone in Section 3 below.

1.2 The Microphone's Role

A microphone is a sensor which transmits the eTube performer's audio signal to the MAs, which enables the audio analysis of that signal, influencing the MA's interaction and output. The Electro-Voice RE20[1] microphone is often used in radio broadcasts. It has a heavy-duty internal pop filter, allowing it to handle the eTube's direct air pressure, and its directionality rejects speaker noise without feeding back. To maintain consistency, our standard setup is an RE20 placed on a mic stand at approximately waist height, although we have tried other mics. For example, we have attached a clip-on mic to the sounding end of the eTube, which maintains a consistent sound input to the mic regardless of the eTube's position or orientation. As a result of this consistent input, the DYCI2 agents would activate too easily and even the softest eTube sounds would launch an output. Therefore, clip-on mics did not allow Davis the possibility to perform solo material without MA interjections. However, an RE20 on a stand allows Davis to move the eTube away from the mic, preventing an MA output.

1.3 Human and Agent Roles

Our research into MAs involves an interdisciplinary team taking on different roles such as programmer, instrument builder, laptop performer, improviser, and composer, among others. The roles are rather fluid, one person may fill multiple functions, or someone may take on a new responsibility. The project's results cannot easily be attributed to one person, but an outcome of the team's combined contributions.

Improvisation is inherently social, and Ingrid Monson reminds us that "in an improvisational situation, it is important to remember that there are always musical personalities interacting, not merely instruments or pitches or rhythms" [23, p. 26]. Suppose one states they are *improvising* with MAs. In that case, this seems to complicate the notion of the social in improvisation as stated by Monson above as we must ask if computers have personalities. Paul Dourish states that interaction between a computer system and end-user is

[1] https://products.electrovoice.com/na/en/re20/

a "fundamentally a social activity" since the computer "mediate[s] communication between the end-user and the system designer" [13, p. 56]. How would the social relationship suggested by Dourish play out in improvisations with MAs? There seems to be a sense that the users interact with the programmer through the constraints and affordances that are built into the software. However, the notion that MAs undertake creative tasks autonomously, as stated in the introduction, also seems to complicate this relationship between developer and user.

We collaborate to create the MA's audio corpora that feed their output, and each team member may take on many of the roles as described below. First, music is improvised and recorded, those recordings are curated for specific material, and then the audio is edited. Then these recordings are trained in the MA software using audio descriptor analysis. This analysis creates a database of segmented and analyzed audio that the MAs use for their output. Finally, in performance, an improviser's live sound is captured by the mic and sent to the computer where it is analyzed using the same audio descriptor analysis system that was used to analyze the corpora. Based on this live analysis, the MA then (re)combines corpora segments in various orders, which are then output via loudspeakers. The corpora recording and training processes thus have a distinct effect on the MA's interaction since they provide the sonic material for the agent, and influence how that material is reorganized and output.

In addition to the audio descriptor analysis, the MAs training is also informed by musical and emotion models. For example, the Musical Agent based on Self-Organizing Maps (MASOM) uses machine listening to analyze various musical features, such as timing structures and harmonic energy filtered to mimic the functions of human perception, in addition to emotional classification using the dimensions of valence and arousal [33]. In other words, when Davis improvises with MAs, he is not simply interacting with (re)produced audio segments. The MA's output is also filtered through machine listening, musical syntax models, and symbolic data embedded in the audio descriptor analysis. Ethical concerns related to biases inherent in these technologies are beyond the scope of this paper.

We consider the MAs as co-creative partners, however, it may not be immediately evident how the MAs contribute to the project or why we might consider MAs as such. Eric Lewis suggests that one may use the intentional stance, make-believe, and hearing persona in music, allowing one to engage with the MAs *as if* they are improvising [20, pp. 57–102]. Following a partial outline of Lewis' approach relevant to this discussion, we will describe certain resultant artistic outcomes in Section 2.

Firstly, Lewis describes Daniel Dennett's intentional stance, which allows one to treat an object like an intentional agent by imagining *as if* it has intentions and beliefs in order to predict its behaviour [10, pp. 13–42]. Lewis then describes how we might engage with MAs emotionally through make-believe, similar to how we engage with fictional characters. Finally, Lewis outlines Jerrold Levinson's "musical personae" theory [19], which describes how "music's expressiveness is a product of a listener imagining that the music is literally expressing emotion" [20, pp. 98–99]. Lewis's argument allows one to consider MAs as an improvising partner regardless of whether or not computers can *actually* improvise and we find this approach compelling for our specific project. Indeed, other researchers who consider MAs as co-creative partners include George E. Lewis, IRCAM's REACH Project,[2] and Notto Thelle and Bernt Wærstad [22, 37].

The team uses different corpora recordings, and we also adjust the interaction settings determined by the developer(s). The MAs' algorithms may be described as a black box, where we may observe the inputs and outputs, but the internal processes are opaque [1, pp. 86–117]. Although the size of the black box may be different to individuals based on their technical knowledge, David Borgo states that the black box requires a "fundamentally performative engagement" [2]. In our project, we learn about the MAs by playing with them. We introduce inputs and observe the outputs, while also adjusting the settings or audio corpora to shape the MA's output further.

2 EMBODIED PERFORMANCE PRACTICE

Davis' eTube performance practice is influenced by the materiality of the eTube, its lightness and flexibility, the vibrations of the low fundamental, and the resultant acoustic sounds when the eTube is moved in space while being played [4]. Indeed, it was first of all this embodied approach to performing with the eTube that informed how the eTube team considered the interaction between Davis and MAs. Below we will discuss the contribution to eTube gestures by other performers.

Developing a new augmented musical instrument (AMI) also involves developing the performance practice surrounding that instrument [17]. The PVC tubing is stored on large spools, which results in the material having a natural curve that influences how the performers hold and move with the instrument. Movement of the eTube and the performer's body has proven to be integral to the instrument's performance practice and the eTube's flexible nature presents intriguing affordances in this respect [11, 13]. This shape has resulted in Davis developing a standard playing position he often uses in performances (see Figure 2).

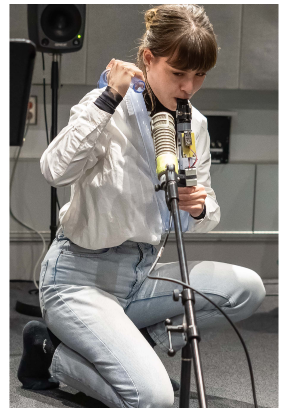
The eTube's design is simple and limited compared to woodwind instruments like the saxophone or clarinet. There are no tone holes to change the pitch, all of the sound adjustments occur via the embouchure and the air stream. Moving the eTube creates phasing effects, spinning it results in the Doppler effect, and partially obstructing the sounding end also filters the sound [12]. Because much of the sound manipulation related to pitch is hidden inside the mouth, the audience does not have the same visual cues or social understanding of the instrumental playing techniques compared with keyed woodwind instruments [9]. Davis also uses contemporary techniques such as slap tongue, multiphonics, and singing while playing, which the public may not have experienced.

Davis' training as a musician and not as a mover is clear as he primarily uses his hands to interact with the eTube and adopts a playing position similar to that of the saxophone [5]. From this standard playing position (see Figure 2), he often rotates his torso back and forth, moving the eTube in a semi-circle around him. In addition, Davis frequently spins the eTube's sounding end with his right hand in various orientations around his body, above his head, and towards the microphone.

[2]http://repmus.ircam.fr/reach

©Vincent Cusson

Figure 3: Bruce playing a tenor saxophone eTube



©Vincent Cusson

Figure 4: Legault playing a bass clarinet eTube

Saxophonist Greg Bruce and clarinetist Maryse Legault have contributed new eTube gestures while co-improvising *3tube*[3] by Pocius as members of the Weather Vane collective, an ensemble dedicated to extended woodwind practice. Davis has been playing with the same eTube for years, and upon buying and building the tenor saxophone and bass clarinet eTubes, we found the new PVC was much stiffer than Davis' existing eTube. This made it less comfortable for Bruce and Legault to use the thumb rest to support the eTube. In addition, Bruce's hands are larger than Davis' so rather than using the thumb rest, Bruce braced his thumb against the bottom of the mouthpiece for support. Due to the stiffer material, the new eTubes were more fatiguing to spin and rotate. In response, Bruce and Legault developed personalized ways to interact with the eTube and mic which were less fatiguing and suited for the 15-minute performance of *3tube*. Notable new gestures include Bruce abruptly pressing the sounding end of the eTube against his body to create a sharp cut-off. In contrast to Davis' approach stated above, Legault would wrap the eTube around her body in various ways while playing. To facilitate Legault orienting and manipulating the eTube on her body, Cusson designed 3D-printed rings to be placed on the eTube. Legault may secure the instrument by hooking her fingers through two different rings while she adjusts her grip on the eTube with her opposite hand or adapting the instrument's position on or around her body. Legault would also move much further from her mic or kneel on the ground to interact with the mic from an intimate proximity (see Figure 4). Inhalation and exhalation remain visible, and perhaps even audible to the public. These compelling performance gestures facilitate interaction with the acoustic space, are tied to the eTube's sonic characteristics, have proven to be intuitive for Davis [8], and have been expanded by gestures established by Bruce and Legault as outlined above.

3 INTERACTING WITH THE MICROPHONE

The microphone could be equated in a simple metaphor as the "ear" of DYCI2 agents since this is how they receive the eTube signal to be analyzed. Put another way, the microphone enables the MA's Umwelt, which refers to how each biological organism maps and understands its own specific way of perceiving the world [38]. Rosemary Lee expands this notion to include a "technological umwelt" that consider the mic as the MA's perceptual apparatus which allows the MA to take in information from its environment [18, p. 2].

What assumptions does one make regarding the microphone when performing with MAs? There seems to be a tacit agreement that the improviser should engage directly with the microphone. If Davis were to move far away from the mic, this would in essence nullify the only "carrier of significance," or the MA's only way to sense and enact agency on its environment [18, p. 1]. If Davis purposefully performs gestures away from the microphone, he will ensure that the mic input is too low to trigger the DYCI2 threshold. In other words, the agent will not "hear" these sounds because the mic input is too low, although these sonic events remain part of the improvisation which may remain audible to other performers and the audience. Although we have spoken about the MAs being collaborators, this is not always an equal collaboration and Davis chooses to assert this control over the MA's umwelt as an artistic affordance. The MA's umwelt comprises fewer modalities than a human, presenting limitations in terms of how Davis can interact with the MAs, compared with another improviser. However, Arne Eigenfeldt and Oliver Bown's Musebots demonstrate how MAs may communicate with each other in performance to synchronize tempo, attacks, pitch sets, or density [14]. Thus, a Musebot ensemble may negotiate change and communicate in its own digital ecosystem, going beyond human communication capabilities.

Davis used certain movements with tubes before the eTube project and other physical gestures have been developed specifically for the eTube mic setup and considering the artistic affordances of the MA's threshold values [31]. Pulling from his experience performing with the saxophone and a variety of microphones, Davis would intuitively move the eTube closer to the mic when he wanted to activate a DYCI2 output. If he played louder material and did not want DYCI2 to respond, he would direct the eTube away from the microphone. This approach allows Davis to shape the MA's output independently of his sonic output by directing the instrument in space relative to the mic.

[3] https://codesdacces.org/evenement/melting-links/

Davis often emphasizes certain movements around the microphone to accentuate triggering the MA's threshold and launching a response. For example, he will often move the eTube towards the mic, and as soon as he performs a slap tongue attack, he quickly pulls the eTube away from the mic. Repeating this gesture also gives a visual emphasis to the attack, which may help audiences anticipate similar sonic gestures. Davis often engages directly with the mic by placing the eTube's sounding end on the RE20's grill. An improviser would not, in general, place their instrument directly on another improviser's face or ears. These ancillary gestures are not directly related to sound production [39] but are movements that interact with the mic, which also might communicate to audiences when Davis is launching an MA's sonic output.

3.1 Other Microphone Interaction Perspectives

It appears that the programmers of these MAs did not intend for the microphone to be interacted with as described above. Thelle's Spire Muse documentation states that audio descriptor analysis is more accurate with a direct signal versus a mic signal [35, p. 180], and later discusses how a contact mic would have been better than a "normal microphone" to record the piano input for Spire Muse [35, p. 234]. Nika describes more generally how "real-time audio, from either a live or prerecorded source" is used for DYCI2's analysis, without specifying microphone details [26, p. 12]. Based on this documentation, the developers refer to a consistent input, with no mention of interacting with the MAs via the microphone as described in this paper.

We question how audiences understand the relationship between improviser, mic, and the MA's output. Mediatized concerts are the norm today, often a fully acoustic concert will be introduced by a person speaking on a mic. In addition, audiences are used to seeing mics on stands or held by vocalists, and for artists to adjust their physical proximity to the mic based on musical or theatrical results [15], for example. Thus, audiences may understand Davis' physical interaction with the microphone and how the eTube's proximity to the mic affects the direct sound received by the microphone, but it may not be clear how these movements interact with threshold levels and as a result the MA's output. The MAs may be appear to audiences as a black box, since the public would observe Davis' input and experience the MA's output, with the MA's internal functions remaining somewhat opaque. How audiences understand the MAs may also be shaped by details we share with them in program notes or spoken before the performance, such as notions of fictional characters, agency, and performance gestures related to the microphone. Each audience member will interpret these details through their own subjective experience depending on their understanding of the technologies and performance techniques, which may affect how they comprehend the human-MA interactions during the performance.

3.2 Embodying Agents Through Spatialization

Referring to Levinson's "musical personae" theory described above, [19], Eric Lewis states that "*all* music listening requires one to take such an imaginative leap, to hear an imaginary persona in the music" [20, p. 98], not only when listening to MAs. Lewis also clarifies that although one may hear a persona in music, one should not assume that those personae reflect an improviser's inner states [20, p. 100]. In other words, we should not assume that if one hears improvised music as "angry" that the improviser is *actually* feeling angry [20, p. 100]. If *all* musical listening involves imagining a persona, then audiences and performers alike take part in a shared, yet subjective listening fiction.

Pocius' interactive spatialization models aim to create complex spatial results from relatively simple interaction [29, 30]. These spatialization systems use pitch and amplitude information extracted from the mic signal to control the placement and movement of the agents throughout the performance. This allows the MAs to adopt movements throughout the loudspeaker system in relation to the acoustic performers' playing style without the need for additional sensors. Michel Chion points out that when a sound source moves, it is much more easily localized compared to a static sound precisely because its location is always changing [6, p. 25]. Pocius' models allow the MAs to mimic and interact in real-time with Davis' performance gestures, providing variety and also cohesiveness to the MA's output through spatialization. And so, if all musical listening involves hearing personae in music, then perhaps by cohesively moving the MA's outputs in the performance space, might this allow the performer and the audiences to imagine the MAs as embodied by the performance space?

4 CONCLUSION

We have presented an embodied approach to performance gestures with the eTube. The eTube is performed with MAs, which present certain challenges to notions of the social and embodiment in improvisation. Following from Eric Lewis' suggestion that one might pretend *as if* the MAs can improvise in order to consider them a co-creative partners, we then demonstrate specific artistic approaches from this stance. These approaches include gestural performance practice based around the microphone, interactive spatialization models eTube's live sound to influence the movement of the MAs through the performance space, perhaps contributing to the sense of the agent being embodied in the space. One might also consider certain ways that the MAs extend the eTube teams' collective agency. In collectively improvised music everyone takes a shared responsibility for the music produced, as described by the core tenet of Tracey Nicholls' "ethos of improvisation" [24, p. 1]. David Borgo describes how group improvisation may result in "complex and emergent properties that are… greater than the sum of its parts" [3, p. 173]. And so, if we are able to engage with MAs in improvisation through make-believe in the same way we engage with a human improviser, then these interactions might also result in complex and emergent properties as suggested by Borgo. Since the eTube team shapes and determines the MA's behaviour to a certain extent, perhaps the MAs are in some sense extending a communal intentionality or agency on behalf of the eTube team members through automatization in performance.

ACKNOWLEDGMENTS

This research is supported by the Social Sciences and Humanities Research Council of Canada (SSHRC), a Natural Sciences and Engineering Research Council of Canada (NSERC) grant supports the Input Devices and Musical Interaction Laboratory (IDMIL) at

McGill, the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), the Metacreation Lab at Simon Fraser University, the International Institute for Critical Studies in Improvisation/Improvisation Festival (IICSI/IF), and the Canada Council for the Arts. Thanks to our supervisors Marie-Chantal Leclair, Robert Hasegawa, Guillaume Bourgogne, and Marcelo M. Wanderley. Thank you to Greg Bruce and Maryse Legault for their artistry and collaboration. Thank you to Eric Lewis for his guidance with this paper. Thank you to Jérome Nika, Sergio Kafejian, Notto Thelle, Ben Carey, and Henning Berg for generously sharing their knowledge of their MAs.

We would like to acknowledge the traditional, ancestral, and unceded Indigenous lands where this project was developed. Tiohtià:ke/Montréal and the surrounding areas have historically been a meeting place for many First Nations. Simon Fraser's Surrey Campus includes the unceded traditional territories of the Semiahmoo, Katzie, Kwikwetlem, Kwantlen, Qayqayt, and Tsawwassen Nations. We strive to respect the history and culture of these diverse communities and to continue to educate ourselves on the impact of our colonial past.

REFERENCES

[1] William Ross Ashby. 1956. *An Introduction to Cybernetics.* Chapman & Hall Ltd., London.

[2] David Borgo. 2016. Openness from Closure: The Puzzle of Interagency in Improvised Music and a Neocybernetic Solution. In *Negotiated Moments: Improvisation, Sound, and Subjectivity*, Gillian Siddall and Ellen Waterman (Eds.). Duke University Press, Durham, NC, USA, 113–130. https://doi.org/10.1515/9780822374497

[3] David Borgo. 2022. *Sync or Swarm, Revised Edition: Improvising Music in a Complex Age* (1 ed.). Bloomsbury Academic, New York. https://doi.org/10.5040/9781501368875

[4] Yves Candau, Jules Françoise, Sarah Fdili Alaoui, and Thecla Schiphorst. 2017. Cultivating kinaesthetic awareness through interaction: Perspectives from somatic practices and embodied cognition. In *Proceedings of the 4th International Conference on Movement and Computing*. 1–8. https://doi.org/10.1145/3077981.3078042

[5] Doga Cavdir and Sofia Dahl. 2022. Performers' Use of Space and Body in Movement Interaction with A Movement-Based Digital Musical Instrument. In *Proceedings of the International Conference on Movement and Computing*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3537972.3537976

[6] Michel Chion. 2015. *Sound: An Acoulogical Treatise.* Duke University Press, Durham. https://doi.org/10.1515/9780822374824

[7] Vincent Cusson and Tommy Davis. 2022. *eTu{d,b}e*: A Preliminary Conduit. In *Proceedings of the International Conference on New Interfaces for Musical Expression.*

[8] Tommy Davis, Kasey Pocius, Vincent Cusson, Philippe Pasquier, and Marcelo M Wanderley. 2023. eTu{d,b}e: Case Studies in Playing with Musical Agents. In *Proceedings of the International Conference on New Interfaces for Musical Expression.*

[9] Kevin Dawe. 2012. The Cultural Study of Musical Instruments. In *The cultural study of music: a critical introduction* (2nd ed.), Martin. Clayton, Trevor. Herbert, and Richard. Middleton (Eds.). Routledge, New York, 195–205. http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S9780203149454

[10] Daniel Clement Dennett. 1987. *The Intentional Stance.* MIT Press, Cambridge, MA, USA.

[11] Balandino Di Donato, Christopher Dewey, and Tychonas Michailidis. 2020. Human-Sound Interaction: Towards a Human-Centred Sonic Interaction Design approach. In *7th International Conference on Movement and Computing*. 1–4. https://doi.org/10.1145/3401956.3404233

[12] Christian Doppler. 1842. On the coloured light of the double stars and certain other stars of the heavens. *Proceedings of the Royal Bohemian Society of Sciences* 2 (1842), 103.

[13] Paul Dourish. 2001. Social Computing. In *Where the action is: the foundations of embodied interaction.* MIT press, Cambridge, MA, USA, 55–98.

[14] Arne Eigenfeldt, Oliver Bown, Andrew R. Brown, and Toby Gifford. 2017. Distributed musical decision-making in an ensemble of musebots: Dramatic changes and endings. In *Proceedings of International Conference on Computational Creativity*. 88–95.

[15] Donna Hewitt. 2003. EMIC-Compositional experiments and real-time mapping issues in performance. In *Converging Technologies: Proceedings of the Australasian Computer Music Association Conference 2003*. Australasian Computer Music Association, 96–104.

[16] Sergio Kafejian. 2023. The Use of an Interactive System in Different Improvisational Contexts. Pre-print, received via personal communication with Tommy Davis. (2023).

[17] Jeff Kaiser. 2018. Improvising Technology, Constructing Virtuosity. *Cuadernos de Música, Artes Visuales y Artes Escénicas* 13, 2 (2018), 87–96.

[18] Rosemary Lee. 2018. The Limits of Algorithmic Perception: technological Umwelt. In *Politics of the Machines-Art and After.* BCS Learning & Development.

[19] Jerrold Levinson. 2006. Musical Expressiveness as Hearability-as-Expression. In *Contemporary Debates in Aesthetics and the Philosophy of Art*, Matthew Kieran (Ed.). Blackwell, Malden, MA, USA.

[20] Eric Lewis. 2019. *Intents and Purposes: Philosophy and the Aesthetics of Improvisation.* University of Michigan Press, Ann Arbor, MI, USA.

[21] George E. Lewis. 2002. Improvised Music after 1950: Afrological and Eurological Perspectives. *Black Music Research Journal* 22 (2002), 215–246. https://doi.org/10.2307/1519950

[22] George E. Lewis. 2021. Co-Creation: Early Steps and Future Prospects. In *Artisticiel / Cyber-Improvisations*, Bernard Lubat, Gérard Assayag, and Marc Chemillier (Eds.). Phonofaune. https://hal.science/hal-03542917

[23] Ingrid Monson. 1996. *Saying Something: Jazz Improvisation and Interaction.* University of Chicago Press, Chicago, IL, USA.

[24] Tracey Nicholls. 2010. Speaking Justice, Performing Reconciliation: Twin Challenges for a Postcolonial Ethics. *Critical Studies in Improvisation/Études critiques en improvisation* 6, 1 (2010), 1–15.

[25] Jérôme Nika, Ken Déguernel, Axel Chemla–Romeu-Santos, Emmanuel Vincent, and Gérard Assayag. 2017. DYCI2 Agents: Merging the "free", "reactive", and "scenario-Based" Music Generation Paradigms. In *Proceedings of the International Computer Music Conference.*

[26] Jérôme Nika, Augustin Muller, Joakim Borg, Gérard Assayag, and Matthew Ostrowski. 2022. *Dicy2 for Max.* Technical Report. Ircam UMR STMS 9912. https://hal.science/hal-03892611

[27] Tone Pernille Østern, Sofia Jusslin, Kristian Nødtvedt Knudsen, Pauliina Maapalo, and Ingrid Bjørkøy. 2023. A Performative Paradigm for Post-Qualitative Inquiry. *Qualitative Research* 23, 2 (April 2023), 272–289.

[28] Philippe Pasquier, Arne Eigenfeldt, Oliver Bown, and Shlomo Dubnov. 2016. An Introduction to Musical Metacreation. *Computers in Entertainment* 14, 2 (2016), 1–14.

[29] Kasey Pocius. 2023. *Expanding Spatialization tools for various DMIs.* Master's thesis. McGill University.

[30] Kasey Pocius, Tommy Davis, and Vincent Cusson. 2024. eTu{d,b}e: Developing and Performing Spatialization Models for Improvising Musical Agents. In *Proceeding of the International Conference on Arts and Humanities*, Vol. 10, no. 1. 20–37. https://doi.org/10.17501/23572744.2023.10102

[31] Jan C. Schacher. 2016. Moving Music: Exploring Movement-to-Sound Relationships. In *Proceedings of the 3rd International Symposium on Movement and Computing*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2948910.2948940

[32] Gillian Siddall and Ellen Waterman. 2016. Introduction: Improvising at the Nexus of Discursive and Material Bodies. In *Negotiated Moments: Improvisation, Sound, and Subjectivity*, Gillian Siddall and Ellen Waterman (Eds.). Duke University Press, Durham, NC, USA. https://doi.org/10.1515/9780822374497

[33] Kıvanç Tatar and Philippe Pasquier. 2017. MASOM: A Musical Agent Architecture Based on Self-Organizing Maps, Affective Computing, and Variable Markov Models. In *Proceedings of the International Workshop on Musical Metacreation*. 1–8.

[34] Kıvanç Tatar and Philippe Pasquier. 2019. Musical Agents: A Typology and State of the Art towards Musical Metacreation. *Journal of New Music Research* 48, 1 (2019), 56–105.

[35] Notto J. W. Thelle. 2022. *Mixed-initiative Music Making. Collective Agency in Interactive Music Systems.* Ph. D. Dissertation.

[36] Notto J. W. Thelle and Philippe Pasquier. 2021. Spire Muse: A Virtual Musical Partner for Creative Brainstorming. In *Proceedings of the International Conference on New Interfaces for Musical Expression.*

[37] Notto J. W. Thelle and Bernt Isak Wærstad. 2023. Co-Creative Spaces: The Machine as a Collaborator. In *Proceedings of the International Conference on New Interfaces for Musical Expression.*

[38] Jakob von Uexküll. 1992. A stroll through the worlds of animals and men: A picture book of invisible worlds. *Semiotica* 89, 4 (1992), 319–391. https://doi.org/10.1515/semi.1992.89.4.319

[39] Marcelo M Wanderley, Bradley W Vines, Neil Middleton, Cory McKay, and Wesley Hatch. 2005. The Musical Significance of Clarinetists' Ancillary Gestures: An Exploration of the Field. *Journal of New Music Research* 34, 1 (2005), 97–113.

[40] Michael Wooldridge. 1999. Intelligent Agents. In *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, Gerhard Weiss (Ed.). MIT Press, Cambridge, MA, USA, 27–72.