The eTube Project: Researching Human-Computer Interaction through an Interdisciplinary Collaboration with Improvising Musical Agents

Tommy Davis

Schulich School of Music
McGill University
Montreal, Quebec, Canada
May 2024

A paper submitted to McGill University in partial fulfillment of the requirements of the degree of D.Mus. Performance Studies

Abstract

This research project outlines the development of a simple augmented wind instrument called the eTube which is played in performance with improvising musical software. An interdisciplinary team including digital luthiers, programmers, composers, and improvisers has contributed to the technological and artistic developments of this project. Rather than creating a new improvising software, existing software developed by other programmers has been combined into a performance framework called *eTu{d,b}e*. As suggested by the word "etude" in this title, and at the heart of this project, is the desire to research the process of learning to work with and adapt interactive technologies, such as improvising software, as an exploration of the research team's artistic and technological vision. The improvised and gestural performance practice with the eTube is the core concept that guides this vision and has led us develop interactive features to build upon the existing software, such as the eTube controller and spatialization models, both of which interact with the improvising software in real-time. Performance case studies will examine the artistic outputs and shed light on the specific approaches to performance practice, interdisciplinary collaboration, technological adaptations, and software updates. Finally, the very notion of improvising computers may seem controversial, and the final chapter presents a philosophical approach for considering the software as a co-collaborative partner and how this stance supports our artistic practice.

Résumé

Ce projet de recherche décrit le développement d'un instrument à vent augmenté simple, appelé eTube, qui est joué en concert avec des logiciels d'improvisations musicales. Une équipe interdisciplinaire composée de luthiers numériques, de programmeurs, de compositeurs et d'improvisateurs a contribué aux développements technologiques et artistiques de ce projet. Plutôt que de créer un nouveau système d'improvisation, des logiciels existants développés par d'autres programmeurs ont été combinés dans un cadre de performance appelé eTu{d,b}e. Comme le suggère le mot « étude » dans ce titre, et au cœur de ce projet, il y a le désir d'étudier le processus d'apprentissage du travail et de l'adaptation des technologies interactives, telles que les logiciels d'improvisation, en tant qu'exploration de la vision artistique et technologique de l'équipe de recherche. La pratique de la performance improvisée et gestuelle avec l'eTube est le concept central qui guide cette vision et qui nous a conduit à développer des fonctions interactives pour développer le logiciel existant, comme le contrôleur eTube et les modèles de spatialisation, qui interagissent tous deux avec le logiciel d'improvisation en temps réel. Des études de cas sur la performance examineront les résultats artistiques et mettront en lumière les approches spécifiques de la pratique de la performance, de la collaboration interdisciplinaire, des adaptations technologiques et des mises à jour logicielles. Finalement, la notion même d'improvisation informatique peut sembler controversée, et le dernier chapitre présente une approche philosophique pour considérer le logiciel comme un partenaire de co-collaboration et comment cette position soutient notre pratique artistique.

Acknowledgements

Thank you first of all to my supervisor Marie-Chantal Leclair for your support, guidance, and inspiration over the past five years. Thank you to my committee members Robert Hasegawa and Guillaume Bourgogne for your detailed thoughts and insights regarding my research and also to the DMus coordinators Simon Aldrich and Lena Weman. Thank you to Professors Marcelo M. Wanderley, Philippe Pasquier, and Eric Lewis for your help with research and papers. A huge thank you to the wonderful collaborators on the eTube project, Vincent Cusson, Kasey Pocius, and Maxwell Gentili-Morin. The eTube project and this document exist thanks to your generosity, artistry, and commitment. Thank you to Jérôme Nika, Sergio Kafejian, Notto Thelle, Ben Carey, and Henning Berg, the programmers and composers who generously shared their knowledge of their improvising systems. Thank you to Greg Bruce and Maryse Legault for playing eTubes with me. Thank you to Nick Zoulek for being an inspiration and always ready to improvise. Thank you to my saxophone teachers Jason Caslor, Glen Gillis, Mark DeJong, William Street, and Jean-Michel Goury.

The eTube has been developed thanks to Student Awards (2021-22, 2022-23, 2023-24) and an Inter-centre Research Exchange Funding from the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT) at McGill University. My research has been supported by a Doctoral Fellowship (2022–24) from 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. Thank you to the Metacreation Lab at Simon Fraser University, the International Institute for Critical Studies in Improvisation (IICSI), the Canada Council for the Arts, and *Conseil des Arts de Montréal*. Thank you to Yves Méthot at CIRMMT and to Richard McKenzie at the Digital Composition Studio (DCS) at McGill for tech and gear support.

Thank you to my parents Casey and Karen and my siblings Jacqui and Robby for your constant love and support. Thank you to my partner Alicja and daughter Rosie for your never-ending love and patience.

Acknowledgements

I 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. I strive to respect the history and culture of these diverse communities and to continue to educate myself on the impact of our colonial past.

Table of Contents

Abstract ..... **ii**

Résumé ..... **iii**

Acknowledgements ..... **iv**

Table of Figures ..... **ix**

List of Acronyms ..... **xi**

Accompanying Media Examples ..... **xii**

Chapter 1: Introduction **1**

1.1 Original Research Contributions and Output 2
1.2 Contextualizing Contemporary Improvisation 9
1.3 Interactive Music, Musical Metacreation, and Computer Creativity 13
1.4 Research-Creation 18

Chapter 2: (e)Tube Context **20**

2.1 Tubes and Duo d'Entre-Deux 22
2.2 First Musical Agent Performances with Saxophones 23
2.2.1 2021 North American Saxophone Alliance Region 10 Conference Online 24
2.2.2 Second Doctoral Recital 25

Chapter 3: Phase 1—Developing the eTube and eTu{d,b}e Framework **28**

3.1 eTube Development 29
3.1.1 Infra-instrument and Constraints 29
3.1.2 eTube Controller and Mouthpiece Adapter Design 31
3.1.3 The "Ear" of the MAs: Microphone Setup 40
3.2 Musical Agents Introduction 41
3.2.1 DYCI2 43
3.2.2 Spire Muse 46

3.2.3 Construction Tools for Interactive Performance: Construction III 49
3.2.4 Algorithmic Comparison 53
3.3 eTu{d,b}e Framework 56
3.3.1 Etude and Naming Scheme 56
3.3.2 Corpora Creation and Curation Process 58
3.3.3 Controller Mapping and Design Process 59
3.3.4 eTube Embodied Performance Practice 61
3.3.5 Performance, Testing, and Artistic Evaluation 64
3.4 Case Studies 66
3.4.1 *live*@CIRMMT 2022 (Cancelled Due to COVID-19) 66
3.4.2 McGill Saxophone Studio Recital Spring 2022 67
3.4.3 New Interfaces for Musical Expression 2022 Conference 68
3.5 Chapter 3 Conclusion 69
Chapter 4: Phase 2—Developing Spatialization Models 71
4.1 Pocius' Spatialization Models 72
4.1.1 The Preset Spatialization Model: DYCI2 and Construction III 73
4.1.2 The Cluster Spatialization Model: Spire Muse 74
4.2 Case Studies 75
4.2.1 Cod'a 2022: *Échanges* 75
4.2.2 *Enfants, apprenez-nous à parler* by Quentin Lauvray 77
4.2.3 Weather Vane: *Improvisation Frameworks* by Greg Bruce and Tommy Davis 79
4.2.4 *live*@CIRMMT: Improvising New Winds 80
4.3.3 North American Saxophone Alliance Biennial Conference 82
4.3.4 Instruments, Interfaces, Infrastructures Conference 84
4.3 Can I Hear You? Monitoring in Multichannel 86
4.4 Chapter 4 Conclusion 87
Chapter 5: Phase 3—Updating the eTube Controller Hardware and Firmware 89
5.1 Resin Mouthpiece Adapters and Key Caps 89
5.2 eTube Controller Updates 92
5.3 Case Studies 95

5.3.1 Weather Vane Collaboration: *3tube* by Kasey Pocius ... 95
5.3.2 Codes d'accès: Melting Links ... 102
5.4 Chapter 5 Conclusion ... 102
Chapter 6: Philosophical Considerations—Do Musical Agents Improvise? ... 104
6.1 Introduction ... 104
6.2 Improvising with Musical Agents ... 105
6.3 The Posthuman, Cyborg, and Media ... 110
6.4 The Intentional Stance ... 115
6.5 Make-Believe and Fiction ... 116
6.6 Embodying Agents Through Spatialization ... 121
6.7 Interacting with the Microphone ... 122
6.8 The Conceptual Nod ... 126
6.9 Ethical Concerns and Care ... 128
6.10 Case Study: Revisiting NIME 2022 ... 132
6.11 Conclusion ... 136
Appendix A ... 141
Appendix B ... 142
Appendix C ... 143
Appendix D ... 144
Appendix E ... 145
Appendix F ... 146
Bibliography ... 147

Table of Figures

Fig. 3.1: eTube alpha version with controller and baritone saxophone mouthpiece 32
Fig. 3.2: eTube beta version with controller and baritone saxophone mouthpiece 32
Fig. 3.3: The author performing the eTube with Spire Muse at Simon Fraser University 33
Fig. 3.4: Above are 3D models of (from left to right) the mouthpiece adapter, battery holder, switch holder, and thumb rest designed by Cusson 35
Fig. 3.5: A 3D-printed plastic mouthpiece adapter for a baritone saxophone mouthpiece 36
Fig. 3.6: The mouthpiece adapters would often break in the same place 36
Fig. 3.7: This baritone saxophone mouthpiece adapter was 3D-printed at a 45-degree angle 37
Fig. 3.8: The overtones accessible by overblowing a 2.54 cm diameter, 219 cm long eTube with a baritone saxophone mouthpiece, transcribed by Quentin Lauvray 38
Fig. 3.9: Some foundational eTube sounds and techniques 39
Fig. 3.10: The DYCI2 interface with three agents including Cusson's added timer bar at the bottom 44
Fig. 3.11: The Spire Muse interface in negotiation mode 48
Fig. 3.12: The Construction III interface with updates by Vincent Cusson 51
Fig. 3.13: The *eTu{d,b}e* framework performance patch including DYCI2 (left) and Construction III (right) with updates by Cusson 54
Fig. 3.14: The *eTu{d,b}e* naming scheme 57
Fig. 3.15: DYCI2 and Construction III mapping table 60
Fig. 3.16: Spire Muse mapping table 60
Fig. 4.1: A workshop to record corpora material and test spatialization models with Pocius at CIRMMT 72
Fig. 4.2: Four spatialization presets from the DYCI2 and Construction III preset model are shown 74
Fig. 4.3: The cluster spatialization model for Spire Muse 75
Fig. 4.4: The first improvised section of *Enfants, apprenez-nous à parler*, staves 3 to 5 78
Fig. 4.5: *Enfants, apprenez-nous à parler*, staves 1 and 2 78
Fig. 4.6: *Enfants, apprenez-nous à parler*, staves 15 and 16 85

Fig. 5.1: Cusson's 3D-printed resin mouthpiece adapter with cork 90
Fig. 5.2: Cusson's 3D-printed resin key caps in two different sizes.. 90
Fig. 5.3: The 3D-printed bass clarinet mouthpiece adapter (left) and the B♭ clarinet mouthpiece adapter (right).. 91
Fig. 5.4: I often broken the eTube wires during transportation 93
Fig. 5.5: The custom-printed PCB (left) and Gentili-Morin's updated controller (right)............. 93
Fig. 5.6: The updated eTube controller designed by Gentili-Morin and Cusson 94
Fig. 5.7: The eTube gamma version with baritone saxophone mouthpiece 94
Fig. 5.8 An excerpt of Pocius' *3tube* score 99
Fig. 5.9: The author's controller mappings for *3tube* 100
Fig. 5.10: Cusson's 3D-printed eTube ring from resin 101

List of Acronyms

AI	Artificial Intelligence
A-Life	Artificial Life
AMI	Augmented Musical Instrument
CCCP	Co-Collaborative Creativity Platform
CSS	Concatenative Sound Synthesis
CTIP	Construction Tools for Interactive Performance
dB	Decibels
DMI	Digital Musical Instrument
DYCI2	Creative Dynamics of Improvised Interaction
ESP	Espressif Designs
IRCAM	Institute for Research and Coordination in Acoustics/Music
JST	Japan Solderless Terminal
MA	Musical Agent
MASOM	Musical Agents based on Self-Organizing Maps
MFCC	Mel-Frequency Cepstral Coefficients
MIDI	Musical Instrument Digital Interface
MuMe	Musical Metacreation
PCB	Printed Circuit Board
PTFE	Polytetrafluoroethylene
PVC	Polyvinyl Chloride
ToM	Theory of Mind
UI	User Interface
VMM	Variable Markov Models

Accompanying Media Examples

This dissertation is being archived and preserved by McGill University in its digital repository, eScholarship, at the following link: www.escholarship.mcgill.ca. To access the digital version and the media files associated with it, use your browser to navigate to the repository and search for the title of this dissertation. The record will list downloadable versions of all digital and media files relevant to this dissertation.

I have created a YouTube playlist containing all the videos for those who wish to stream the videos rather than downloading them from eScholarship. This playlist link will be cited in the footnotes throughout this dissertation and is accessible here: https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg. In addition, you may access these videos and the YouTube playlist via my website at www.tommydavis.ca/dissertation-media. The titles of the videos in the YouTube playlist should correspond with the reference cited throughout. I anticipate that due to technological or human error, these links will cease to function in the coming years. In principle, the eScholarship page is a permanent storage location, which should remain accessible well into the future.

Chapter 1: Introduction

For the past five years I have been researching the eTube and the *eTu{d,b}e* improvisation framework during my doctoral studies at McGill University. The eTube is a modest wind instrument made from a baritone saxophone mouthpiece and plastic polyvinyl chloride (PVC) tube. It is augmented with a wireless controller to facilitate interaction with improvisation software in musical performance. The *eTu{d,b}e* framework describes the various improvisation software that has been adopted and adapted by a collaborative interdisciplinary team including digital luthiers, programmers, composers, improvisers, and scholars. This is the software that I improvise with, and the eTube controller allows me to interact with it wirelessly during performance. Through this project I hope to present specific approaches for interacting and improvising with improvising software, the process of learning to use and perform with software, the various roles of collaborators in this process, specific performances that advance artistic, technological, and a philosophical position supporting this approach to improvising software.

I have had extensive formal musical training in saxophone performance, including interpreting works including electronic parts. However, prior to my doctorate, I had no formal training in electronic music composition, music production, or recording and my skills in this domain have been primarily self-taught. As a performer-improviser who has been involved in electronic music for years, I have depended on others' support for endeavors outside my skillset. I have been privileged to dedicate five years to my doctoral studies, more than three of those years to improvising agents, and the last three years specifically to the eTube and *eTu{d,b}e* project. This project has required significant time, funding, gear, expertise, and artistic and technological resources. This was precisely why I began a doctorate. On my own, I could not devote the necessary time, did not have the skills, and could not properly remunerate collaborators to successfully undertake certain projects I had in mind. Although my doctoral research changed significantly since my initial project proposal, the core aspects including live electronic music, contemporary saxophone techniques and noise, and improvisation have maintained consistent throughout. Reflecting on the time, funding, and human and technological resources that have supported and facilitated this research, I could not have developed this project outside of a research institution.

This paper first outlines a historical context for this research, critical scholarly and artistic influences, and early developments in my research at McGill, contextualizing the project and framing the original contributions of this research. I will then outline the project developments from 2021 to 2024 that were supported by three Student Awards from the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), undertaken with music technology master's students Vincent Cusson, Kasey Pocius, and Maxwell Gentili-Morin.[1] This will include a discussion of the specific collaborations between myself, as a saxophonist, performer, and improviser, with each member of the collaborative team. Finally, I will discuss certain philosophical arguments to situate this research within a larger social and political context including issues of fictional persona, artificial intelligence, cyborg theory, and caring for imaginary entities.

1.1 Original Research Contributions and Output

There are many artist-developers who have designed and perform with their own improvising software such as Sandeep Bhagwati's Native Alien, Benjamin Carey's _derivations, George E. Lewis' Voyager, Robert Rowe's Cypher, and Michael Young's NN Music.[2] Although these

---

[1] "Welcome to CIRMMT," CIRMMT, accessed March 3, 2024, https://www.cirmmt.org/en.

[2] Sandeep Bhagwati, "Virtuosities of the Native Alien," in *Contemporary Musical Virtuosities*, ed. Louise Devenish and Cat Hope, Routledge Research in Music (London: Routledge, 2024), 68–78, https://doi.org/10.4324/9781003307969; Benjamin Carey, "_derivations: Improvisation for Tenor Saxophone and Interactive Performance System," in *Proceedings of the 9th ACM Conference on Creativity & Cognition* (Sydney, Australia: ACM, 2013), 411–12, https://doi.org/10.1145/2466627.2481226; George E. Lewis, "Too Many Notes: Computers, Complexity and Culture in 'Voyager,'" *Leonardo Music Journal* 10 (2000): 33–39, https://www.jstor.org/stable/1513376; Robert Rowe, "Machine Listening and Composing with Cypher," *Computer Music Journal* 16, no. 1 (1992): 43–63, https://doi.org/10.2307/3680494; Michael Young, "NN Music: Improvising with a `Living' Computer," in *Computer Music Modeling and Retrieval. Sense of Sounds*, ed. Richard Kronland-Martinet, Sølvi Ystad, and

artists collaborate with other improvisers to perform their software, the developers are most often present during performances and are involved with certain elements of the performance. There are few research documents investigating improvisers who work with existing software and the process of collaboratively adapting and learning to perform with improvising software for one's needs. Indeed, the issue of other users in the performance of new digital interfaces, which includes improvising software, is a challenge for longevity and dissemination of these instruments or interfaces.[3] Projects are often developed and performed by their creators only to be abandoned a few years later, often with no adoption by other users.[4] As I will discuss below, collaborating with composers on new works and finding other artists to perform with the eTube has been an important goal for this project to ensure its dissemination and longevity (see Sections 4.2 and 5.3). The eTube project is one example of how an interdisciplinary team learns to work with existing improvising software created by other programmers, and the process of adapting these systems for our own creative interests, rather than developing our own improvising software.

This project has taken on a life of its own, with each new collaborator bringing their own ideas and technological and artistic direction to the project. At times this has been a clear process regarding who is responsible for the artistic directions or improvements, and at other times, this has been a complicated jumble of influences which resulted in new discoveries that we would not have made on our own. Throughout this document I will attribute credit to each collaborator,

---

Kristoffer Jensen (Berlin: Springer, 2008), 337–50, https://doi.org/10.1007/978-3-540-85035-9_23.

[3] Fabio Morreale and Andrew McPherson, "Design for Longevity: Ongoing Use of Instruments from NIME 2010-14," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Aalborg University Copenhagen, Denmark, 2017); Raul Masu, Fabio Morreale, and Alexander Refsum Jensenius, "The O in NIME: Reflecting on the Importance of Reusing and Repurposing Old Musical Instruments," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Mexico City, Mexico, 2023).

[4] Clayton Rosa Mamedes et al., "Composing for DMIs - Entoa, Music for Intonaspacio," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Goldsmiths, University of London, UK, 2014), 509–12; Morreale and McPherson, "Design for Longevity."

while also underlining the intertwined situations where many people have contributed to the project, including aesthetic, hardware, software, performative, philosophical, conceptual, design, and artistic elements.

I can say that few details in this project can explicitly be attributed only to me, although the overall project has inevitably been guided in certain directions by my own interests and artistic preferences. However, I also acknowledge that this type of collaborative research, which I find so fulfilling, requires that the collaborators be willing and open to share in a messy and iterative co-creative process. The project's outcomes are a result of the layered interactions on musical, social, interpersonal, and technological levels which reflect the specific influences and expertise of both human and non-human actors involved. It is because of these individuals' dedication that the project has been cultivated into its current representation here.

As a technology-based artistic project, it is in a constant state of updates and upgrades, with representative performances as snapshots of the project at a given time, while the development process continues. *eTu{d,b}e* is also an ongoing professional endeavour, which we anticipate will continue following this dissertation, and as such, the details outlined here are current as of March 2024. As will be discussed below, the focus will be on the process that has brought us to this point, and which have inspired the philosophical questions that seem pertinent to us at this time. I hope that this document will prove useful and inspiring for others in similar, or distinctly different projects, and will suggest what might be possible via various technological approaches and applications as well as what questions one might ask throughout the process.

The following list summarizes the outputs for this project which have been realized over the course of my studies:

1. Performance Media, Commissioned Works, and Educational Outreach

Performance media and educational outreach material have been disseminated through publications and online.

A. Performance Media

Cusson, Vincent, and Tommy Davis. "Etu{d,b}e: A Preliminary Conduit." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2022. https://doi.org/10.21428/92fbeb44.c05957ee.

eTube – Spire Muse Sessions | Take 2. YouTube video, 3:11, 2022. https://www.youtube.com/watch?v=49LuS84ZOxw.

eTube – Spire Muse Sessions | Take 5. YouTube video, 3:57, 2022. https://www.youtube.com/watch?v=Rq5HJ07etOI.

eTube – Spire Muse Sessions | Take 6. YouTube video, 5:41, 2022. https://www.youtube.com/watch?v=nbHG0cFNafs.

B. Education and Outreach

Davis, Tommy. "Computer Agency and Improvisation". YouTube video, 07:37, May 2024. https://www.youtube.com/@resonatortube4567. [educational video]

Mar. 20, 2024, ANTH 555 – Sonic Ethnography: Theory and Practice, McGill, Montreal. "eTu{d,b}e: embodying improvising musical agents through a spatialized eTube practice," with Kasey Pocius and Vincent Cusson [guest presentation].

Apr. 11, 2023, McGill Association for Student Composers (MASC) Presentation, Digital Composition Studio, McGill University. "eTu{d,b}e: Improvising with Musical Agents and the eTube," with Vincent Cusson [guest presentation].

Feb. 23, 2023, MUS 3323 Musique de création et technologies, Université de Montréal, Montreal. "eTu{d,b}e: Improvising with Musical Agents and the eTube," with Vincent Cusson [guest presentation].

2. Commissioned Works

Three works have been commissioned for the eTube.

Bruce, Greg, and Tommy Davis. *Improvisation Frameworks*. 2023.

Pocius, Kasey. *3tube*. 2023.

Lauvray, Quentin. *Enfants, apprenez-nous à parler*. 2022.

3. Published Research Papers

I have co-authored three research papers on the eTube and *eTu{d,b}e* framework, including one forthcoming book chapter.

Forthcoming: Davis, Tommy, and Vincent Cusson. "Creative Cyborgs: Researching Human-Computer Interaction with the eTu{d,b}e Improvisation Framework." In *Music and Transcendence in a Posthuman Age*, edited by Ariane Couture, Zoey Cochran, and Kit Soden, 2025.

Davis, Tommy, Kasey Pocius, Vincent Cusson, Maxwell Gentili-Morin, and Philippe Pasquier. "Embodied eTube Gestures and Agency." In *Proceedings of the 9th International Conference on Movement and Computing*. New York, NY: Association for Computing Machinery, 2024. https://doi.org/10. 1145/3658852.3659084.

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

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

4. Conference Presentations, Concerts, and Workshops

I have presented, co-presented, and performed this project locally and internationally for concert promoters, festivals, conferences, workshops, and university seminars. Performances are one of the primary cultural artifacts created during this project. We have participated in workshops by distinguished artists and researchers working with interactive technology.

A. Conference Presentations

Forthcoming: May 29–Jun. 2, 2024, The 9th International Conference on Movement and Computing, Utrecht, The Netherlands. "Embodied eTube Gestures and Agency" [poster presentation].

Apr. 8–13, 2024, Sonorities Festival, Queen's University, Belfast, Northern Ireland. Symposium Talks: "eTu{d,b}e: Exploring Musical Agents through Improvisations with an Infra-Instrument," with Kasey Pocius [presentation]. Handmade Music: eTu{d,b}e with Spatialised Improvising Agents (UK Premiere) by Kasey Pocius and Tommy Davis [performance].

Oct. 20–22, 2023, Sound, Meaning, Education: CONVERSATIONS & *improvisation*s, International Institute for Critical Studies in Improvisation (IICSI), University of Guelph, Ontario. "eTu{d,b}e*:* Embodying Improvising Musical Agents through a Spatialized eTube Practice," with Kasey Pocius [lecture-recital].

Sep. 8, 2023, International Conference on Arts and Humanities (ICOAH), Bangkok, Thailand (online). "eTu{d,b}e: Further Case Studies into Instrument Design and Playing with Spatialized Musical Agents," with Kasey Pocius [presentation, second author].

May 31–June 2, 2023, New Interfaces for Musical Expression (NIME) Conference, Mexico City, Mexico. "eTu{d,b}e: Case Studies in Playing with Musical Agents," with Kasey Pocius [presentation].

May 26, 2023, CIRMMT, OICRM, BRAMS Student Colloquium (COBS), UQÀM, Montreal. "eTube: Improving the Physical Interface for Reliability in Performance and Adding Communication Between Performer and Associated Musical Agents," with Maxwell Gentili-Morin [presentation]. "Utilizing eTube Performance Gestures to Spatialize Autonomous Musical Agents Using Various Spatialization Models," with Kasey Pocius [presentation].

May 11–13, 2023, Harvard University Instruments, Interfaces, Infrastructures: An Interdisciplinary Conference on Musical Media Conference, Harvard University, Cambridge, MA. eTu{d,b}e with Apatialised Agents by Kasey Pocius and Tommy Davis [performance]. Enfants, apprenez-nous à parler (US premiere) for eTube by Quentin Lauvray, sptialisation by Kasey Pocius [performance].

Apr. 2, 2023, North American Saxophone Alliance (NASA) Biennial Conference, University of Southern Mississippi, Hattiesburg, MS**.** *eTu{d,b}e* (US premiere) with Spatialised Agents by Kasey Pocius and Tommy Davis [performance].

Mar. 25, 2023, Doctoral Lecture-Recital, Tanna Schulich Hall, McGill University. "Improvising Cyborgs: Researching Computer Creativity with the *eTu{d,b}e* Framework."

Feb. 25, 2023**,** Music and Transcendence in a Posthuman Age International Conference, Centre Pierre-Péladeau, Montreal. "Improvising Cyborgs: Researching Computer Creativity with the *eTu{d,b}e* Framework," with Vincent Cusson and Quentin Lauvray [lecture-recital].

Feb. 23, 2023, Musical Creation and Technologies Seminar, Université de Montréal. "eTu{d,b}e: Improvising with Musical Agents and the eTube," with Vincent Cusson [guest presentation].

May 24, 2022, CIRMMT, OICRM, BRAMS Student Colloquium (COBS), McGill, Montreal. "Designing a Physical Interface to Facilitate Interaction with an Autonomous Musical Agent in Improvised Performance," with Vincent Cusson [presentation]. "Utilizing eTube Performance Gestures to Spatialize Autonomous Musical Agents Using Various Spatialization Models," with Kasey Pocius [presentation].

B. Concerts

Nov. 30, 2023, melting links presented by Codes d'accès, CIRMMT Multimedia Room, Montreal. *3tube* for baritone, tenor & clarinet eTubes, agents, and electronics by Kasey Pocius, Tommy Davis, Maryse Legault, and Greg Bruce.

Dec. 9, 2022, Improvising New Winds presented by *live*@CIRMMT, CIRMMT Multimedia Room, Montreal. *Enfants, apprenez-nous à parler* (premiere) by Quentin Lauvray. *Improvisation Frameworks* (premiere) by Greg Bruce and Tommy Davis.

Oct. 18, 2022, Cod'a 2022: *Échanges* presented by Codes d'accès, Eastern Bloc, Montreal. *eTu{d,b}e*, improvisation libre pour eTube by Kasey Pocius, Tommy Davis, and Vincent Cusson.

C. Workshops

Dec. 1, 2023, Composing and Performing with Digital Musical Instruments, CIRMMT RA4 Workshop, McGill University, Montreal. "*3tube*: co-composing for an infra-instrument, fixed media, and agent software," with Kasey Pocius and Maryse Legault [presentation and workshop].

Oct. 17, 2023, Exploring machine learning, artificial creativity and human musicality, with Marc Chemillier, McGill, Montreal. *eTu{d,b}e* with spatialised improvising agents, with Kasey Pocius [presentation and performance].

Jan. 25, 2023, Workshop in Digital Musical Instruments with Pamela Z, McGill, Montreal. eTu{d,b}e: improvising with musical agents and the eTube [presentation and performance].

Apr. 18–May 16, 2022, CIRMMT Inter-centre Research Exchange with Prof. Philippe Pasquier, Metacreation Lab, Simon Fraser University, Vancouver, Canada.

5. Development of a New Augmented Instrument and Performance Practice

The eTube instrument and integrated the controller is a newly invented instrument. The microphone setup, performance gestures, and techniques have been developed considering this instrument's technical constraints, and the performance practices afforded by the improvising software.

6. Development of Software for Performance

The *eTu{d,b}e* framework is made of existing software developed by other programmers and we have combined these in novel ways. In addition, the eTube team members have developed new software or modules specific to our performance practice, which may be adapted for other contexts.

7. Collaborative Process with Technologists

Discussion of our collaboration working with improvising software and electronic media, the different roles that members played throughout, and our proposition for creatively performing with this software. The specific technologists I work with are also trained musicians, however, their approach as programmers and developers has challenged me to adopt a more reflexive and critical approach to the creation process, not only a focus solely on the final artistic product, which has manifested in the research documents presented here.

8. Philosophical discussion

Throughout this project, including previously published papers, I describe an approach to working with improvising software inspired by philosophical literature on machine improvisation, fiction, and posthuman theory. This discussion is specific to the current project, but also speaks to larger questions about human-human interaction, human-machine interaction, improvisation, collaborative creation processes, and gestural performance practice.

1.2 Contextualizing Contemporary Improvisation

In a landmark book on the subject, Derek Bailey stated that "improvisation enjoys the curious distinction of being both the most widely practiced of all musical activities and the least acknowledged and understood."[5] When we first begin talking about improvisation, we must ask ourselves where we are situated and in what community are we referring to since improvisation practice varies between geographical locations and communities. Indeed, George E. Lewis defines improvised music 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."[6] In addition to the social aspect, improvised performance is a highly embodied practice.[7] 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. The focus that the social and embodied occupy in improvisation challenges those working with improvising software since one must consider how the machines enter this social location as disembodied entities. This forces one to think more deeply about what the social is, and what is important about bodies in improvisation.

In foundational improvisation research, Jeff Pressing proposed the term "referent" to refer to "a set of cognitive, perceptual, or emotional structures (constraints) that guide and aid in the production of musical material."[8] As Pressing explains, the referent may refer to a musical

---

[5] Derek Bailey, *Improvisation: Its Nature and Practice in Music* (New York: Da Capo Press, 1993), ix.

[6] George E. Lewis, "Improvised Music after 1950: Afrological and Eurological Perspectives," in *The Other Side of Nowhere: Jazz, Improvisation, and Communities in Dialogue*, ed. Daniel Fischlin and Ajay Heble, 1st ed., Music/Culture (Middletown, CT: Wesleyan University Press, 2004), 149, http://catdir.loc.gov/catdir/toc/ecip048/2003019192.html.

[7] Gillian Siddall and Ellen Waterman, "Introduction: Improvising at the Nexus of Discursive and Material Bodies," in *Negotiated Moments: Improvisation, Sound, and Subjectivity*, ed. Gillian Siddall and Ellen Waterman, Improvisation, Community, and Social Practice (Durham: Duke University Press, 2016), 1, https://doi.org/10.1515/9780822374497.

[8] Jeff Pressing, "Psychological Constraints on Improvisational Expertise and Communication," in *In the Course of Performance: Studies in the World of Musical Improvisation*, ed. Bruno Nettl and Melinda Russell, Chicago Studies in Ethnomusicology (Chicago, IL: University of Chicago Press, 1998), 52.

structure, such as a song form, which helps improvisers prepare variations on material, minimizing processing capacity, decision-making, and anxiety during performances in addition to providing shared references between performers and increasing the possibility of synergistic alignment. A referent also allows an improviser to prepare material ahead of time, to anticipate developments in the moment.

Bailey has put forth the notion of "non-idiomatic" where the artists are not expressing a specific idiom, such as jazz or flamenco music.[9] In contrast to the referent above, an improviser would not prepare for specific developments. Rather, this is negotiated between improvisers during the improvisation. Framing improvisation as non-idiomatic suggests a kind of ideal to strive towards, which avoids referencing established idiom. Some may argue that the non-idiomatic also becomes an idiom after a certain time. Non-idiomatic improvisation is indeed an approach adopted by many improvisers; however, the definition suggests a limited case scenario, and does not consider the differing practices across communities as indicated by Lewis above. Improvisation is distinct from other artistic processes like composition and musical interpretation, which are primarily constructs from a Eurological musicking perspective. Defining improvisation in relation to these practices, as in "real-time composition" is also problematic as it defines improvisation within a Eurological framework.[10] Indeed, as David Borgo states that "referent-free improvisation involves the development of a personal, enculturated knowledge base as much as any other practice. In fact, developing an identifiable voice or an individual/ensemble style is an essential part of establishing one's expertise."[11] Improvisation is informed by learned gestures, patterns, and styles which are develop over time by individuals or groups.

As suggested above, this situates improvisation as a corporeal practice, involving bodies and developed with and through bodies. This corporeal practice is problematized when one considers

---

[9] Bailey, *Improvisation*, xi–xii.

[10] Lewis, "Improvised Music After 1950," 152–53.

[11] David Borgo, *Sync or Swarm, Revised Edition: Improvising Music in a Complex Age*, 1st ed. (New York: Bloomsbury Academic, 2022), 27, https://doi.org/10.5040/9781501368875.

how one might improvise with disembodied software that exists on a computer and is diffused through loudspeakers. Throughout this document, I hope to demonstrate ways our interdisciplinary team has developed an expertise working with existing improvising software and the resultant artistic outputs. In addition, this project aims to show how improvisation with musical agents as a generative process which informs new ways of embodied improvisation practice, artifact creation, and knowledge. As such, this document will describe one specific approach to improvisation which is situated within a global community of improvisers who each maintain their own cultural and musical practices.

Throughout this text I will use the terms improvisation and improvised music to refer to my improvisation practice which is situated within a specific geographical, cultural, and social community, as suggested by Lewis above. My education has been primarily in Eurological musicking involving interpreting scores and performance in concert venues designed specifically for this practice. I have been involved in the Montreal improvisation scene, which has had a significant influence on my improvisation practice over the past ten years. As a white cis-gender man, my relative privilege within society results in advantages and means that I may hold certain biases based on this positionality. Through this project, I am interested in questioning hierarchies often assumed in collaborative Eurological musicking, such as composer and performer roles and the relationship between improvisation and scores. I am also interested in reflecting on my own improvisation practice through constraints imposed by an instrument with limited pitch virtuosity, and the creative possibilities afforded by these limitations. After describing the projects evolution over the past three years, I will engage with philosophical literature in the final chapter to support my initial assumptions that one may indeed improvise with software. Even if we cannot objectively determine if the software is improvising, by pretending *as if* the software is improvising, we can engage in compelling performances, nonetheless. I will also argue that the agency we attribute to the software acts in part to extend the collective agency of the collaborators involved in the project. In this regard, one of the challenges has been to present how this work is undertaken by an interdisciplinary team that shares in the conception, creation, production, and performance aspects, while acknowledging the intertwined contributions of collaborators and attributing credit to individuals where merited. Finally, as an extension of my own agency, the improvising software also presents the possibility to reflect on my own improvisation practice, what I value in improvising partners, and my own biases that shape these relationships.

1.3 Interactive Music, Musical Metacreation, and Computer Creativity

Human-computer interaction (HCI) is a multidisciplinary field which studies "the way in which computer technology influences human work and activities."[12] Within the broad field of HCI are foundational approaches to music making with machines. This includes interactive music systems which are described by Robert Rowe as "those whose behavior changes in response to musical inputs."[13] Rowe distinguishes between two approaches using existing musical terminology. The "instrument" paradigm is a system that extends a musical instrument's capabilities, and the "player" paradigm is a system that tries to construct an artificial performer to play with a human.[14] In addition, Todd Winkler has suggested four models for interactive performance on established practices which he calls the "conductor," "chamber music," "jazz combo improvisation," and "free improvisation" models.[15] The conductor model is based on pre-determined music and allows a performer to influence temporal aspects of the piece in real-time. The chamber music model is based on written scores where there is reciprocal influence between the performers and the electronics. The jazz combo model is based on shared assumptions and implied rules from years of collective experience which are coded into the electronics to allow interaction with a performer within the specified context or aesthetic. The free improvisation

---

[12] Alan Dix, "Human-Computer Interaction," in *Encyclopedia of Database Systems*, ed. Ling Liu and M. Tamer Özsu (Boston, MA: Springer US, 2009), 1327, https://doi.org/10.1007/978-0-387-39940-9_192.

[13] Robert Rowe, *Interactive Music Systems: Machine Listening and Composing* (Cambridge, MA: MIT Press, 1993), 1.

[14] Rowe, 8.

[15] Todd Winkler, *Composing Interactive Music: Techniques and Ideas Using Max* (Cambridge, MA: MIT Press, 1998), 23–27, https://proxy.library.mcgill.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1430&scope=site.

model is highly interactive where human and computer influence each other, and the computer's output is "listenable music on its own" (see Section 6.5).[16] The notions of control and agency in these models becomes less specific as one traverses the categories from conductor to free improvisation. This project focuses mainly on improvising software that aligns primarily with Rowe's "player" and Winkler's "free improvisation" model, although there are aspects of the "instrument" paradigm as well.

Until this point, I have been using the words "improvisation software" to refer to the computer programs that we work and perform with. I would like to propose a more specific terminology that I will continue to use throughout this document. Researchers often describe improvising software using anthropomorphisms and the term "musical agent" (MA), which is 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" in real-time.[17] We also use the term musical agent, and although we strive to limit anthropomorphisms so as not to overstate the MAs' abilities (see Section 6.9), anthropomorphizations are often a useful strategy to convey complex details in an accessible format. Michael Wooldridge states "an *agent* is a computer system that is *situated* in some *environment*, and that is capable of *autonomous action* in this environment in order to meet its design objectives" (emphasis in original).[18] In addition, MAs "explore the notions of autonomy, reactivity, proactivity, adaptability, coordination and emergence"[19] in non-deterministic environments, where the same action may result in different effects.[20] We distinguish between

[16] Winkler, 26.

[17] Kıvanç Tatar and Philippe Pasquier, "Musical Agents: A Typology and State of the Art towards Musical Metacreation," *Journal of New Music Research* 48, no. 1 (January 1, 2019): 56, https://doi.org/10.1080/09298215.2018.1511736.

[18] Michael Wooldridge, "Intelligent Agents," in *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, ed. Gerhard Weiss (Cambridge, MA: MIT Press, 1999), 15, http://cognet.mit.edu/library/books/view?isbn=0262731312.

[19] Tatar and Pasquier, "Musical Agents," 62.

[20] Wooldridge, "Intelligent Agents," 16.

objects and agents, where agents are autonomous and proactive, meaning that they "exhibit goal-directed behaviour by *taking the initiative* in order to satisfy their design objectives" (emphasis in original).[21] This means that once started, MAs will operate autonomously by generating novel musical output in real-time based on some kind of stimulus. Devices like Mp3 players are not MAs since they are not proactive; when music is played back the content is simply (re)produced. MA research is part of musical metacreation (MuMe), which Philippe Pasquier defines 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."[22] MuMe uses the tools developed by artificial intelligence (AI) and machine learning research which allows artists to design autonomous systems to undertake various musical tasks.[23] Whereas AI research focuses on solving problems with optimal solution, in MuMe, like in art and improvisation, there are no "optimal solutions," and artists must define their own constraints, subjective artistic goals, and outcomes.[24]

Kıvanç Tatar and Philippe Pasquier define a spectrum of MA autonomy on a continuum "ranging from purely reactive systems without autonomy to completely autonomous systems."[25] In addition, they define a continuum "that ranges from specific systems to purely generic systems."[26] The reactive-autonomous spectrum deals with reactive agents that are more like guitar effects pedals compared to autonomous agents, which are proactive. The specific-generic spectrum describes specific agents which are designed for a certain aesthetic and performance situation, compared with generic agents which may learn new material in real-time, or may

---

[21] Wooldridge, 23.

[22] Philippe Pasquier et al., "An Introduction to Musical Metacreation," *Computers in Entertainment* 14, no. 2 (December 2016): 2:4, https://dl.acm.org/doi/10.1145/2930672.

[23] Arne Eigenfeldt et al., "Towards a Taxonomy of Musical Metacreation: Reflections on the First Musical Metacreation Weekend," *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* 9, no. 1 (2013).

[24] Pasquier et al., "An Introduction to Musical Metacreation," 2:2.

[25] Tatar and Pasquier, "Musical Agents," 61.

[26] Tatar and Pasquier, 61.

function in diverse performance aesthetics. In Section 3.2 below I will describe the various agents we use, which all fall along these spectra.

According to Geraint Wiggins, "the zenith of human intelligence is very often portrayed as the ability to create, and to create radically new and/or surprising things."[27] Wiggins goes on to propose a working definition of creativity for machines which is described as "the performance of tasks which, if performed by a human, would be deemed creative." In the case of musical agents, we primarily judge the agents based the musical context, and as a result, their sonic output. In performance, the agents may appear *as if* they are improvising and interacting in a musical way. However, we also know that computers behave as humans have programmed them to behave. As mentioned above, agents are proactive, and they may operate in non-deterministic ways, which can result in surprising audio output. An answer to the larger question of whether computers can actually improvise is beyond the scope of this document, however, in Chapter 6 I will discuss a philosophical stance that allows us to interact with computers in improvised performance without needing to answer that question.

Using machines to create music dates back centuries.[28] However, recent advances in AI, its increased visibility in the media, and the integration of these tools in our workplaces and lives has resulted in polarizing perspectives on these technologies. Within this continuum of perspectives and using the tools of MuMe, there are researchers and artists who have adopted the perspective that one may consider MAs and the associated technologies as co-collaborative partners.[29] Indeed, many researchers working in MuMe would consider MAs as a "partner,"

---

[27] Geraint A. Wiggins, "A Preliminary Framework for Description, Analysis and Comparison of Creative Systems," *Knowledge-Based Systems* 19, no. 7 (2006): 450, https://doi.org/10.1016/j.knosys.2006.04.009.

[28] Teun Koetsier, "On the Prehistory of Programmable Machines: Musical Automata, Looms, Calculators," *Mechanism and Machine Theory* 36, no. 5 (May 1, 2001): 589–603, https://doi.org/10.1016/S0094-114X(01)00005-2.

[29] Notto J. W. Thelle and Bernt Isak Wærstad, "Co-Creative Spaces: The Machine as Collaborator," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Mexico City, Mexico, 2023).

which is defined as “a person who takes part with another or others in doing something.”[30] In contrast, a collaborator is defined as “one who works in conjunction with another or others.”[31] Whereas one might agree that a MA is taking part in an improvisation simply by its presence in the moment, the agent as a collaborator may be more of a stretch. Working in conjunction with someone seems to imply that two parties are working towards a shared goal. Indeed, other researchers who consider MAs as co-collaborative partners include Gérard Assayag and colleagues at the *Institut de recherche et coordination acoustique/musique*’s (IRCAM) Music Representations Team, George Lewis’ Voyager, and Notto Thelle and Burnt Wærstad’s Co-Creative Communication Platform (CCCP), among others.[32] The eTube project is one among many that considers the various ways MAs might be considered as co-collaborative partners in improvised performance. Co-collaboration suggests that MAs work in conjunction with others, and so one must then ask if the agents are *intentionally* collaborating with an improviser? Considering the agents as co-collaborators is problematic in certain ways, including the question of intention. In Chapter 6, I will discuss notions such as the intentional stance, fictional characters, make-believe, and ethical reasons to explain how and why I have chosen to consider the MAs as co-collaborators in specific situations.

Computer processing relies on syntax for programs to be compiled or interpreted. Syntax is defined as “the set of rules and principles in a language according to which words, phrases, and clauses are arranged to create well-formed sentences.”[33] Syntax deals with structural details but

[30] “About,” Musical Metacreation, November 12, 2015, https://musicalmetacreation.org/about/; Oxford English Dictionary, “Partner, n.1, Sense 2.a” (Oxford University Press, December 2023), Oxford English Dictionary, https://doi.org/10.1093/OED/1044963391.

[31] Oxford English Dictionary, “Collaborator, n., Sense 1” (Oxford University Press, July 2023), Oxford English Dictionary, https://doi.org/10.1093/OED/5793806289.

[32] “Music Representations Team,” accessed February 24, 2024, http://repmus.ircam.fr/impro; George E. Lewis, “Co-Creation: Early Steps and Future Prospects,” in *Artisticiel / Cyber-Improvisations*, ed. Bernard Lubat, Gérard Assayag, and Marc Chemillier, Dialogiques d’Uzeste (Phonofaune, 2021), https://hal.science/hal-03542917; Thelle and Wærstad, “Co-Creative Spaces.”

[33] Oxford English Dictionary, “Syntax” (Oxford University Press, September 2023), Oxford English Dictionary, https://doi.org/10.1093/OED/1187492139.

does not relate to meaning. Semantics is "the branch of linguistics or philosophy concerned with meaning in language."[34] Although computer processing indeed has syntax and developments in artificial intelligence (AI) have significantly advanced computer capabilities, there are strong arguments against whether computers can understand meaning, such as John Searle's essay known as the Chinese Room Argument.[35] Musical syntax is often well defined; performance practice, composition, and improvisation are often governed by specific rules and conventions. However, semantics, or musical meaning, is often rather subjective. The subjective meaning that one ascribes to music through interpretation and analysis depends on the individual person as well as cultural factors. This has been a brief introduction to these concepts and issues, which will be discussed throughout the text and in greater depth in Chapter 6.

1.4 Research-Creation

Research-Creation is a relatively recent interdisciplinary approach that combines academic research with creative practices. The methodology emphasizes the integration of research and creative expression, blurring the traditional boundaries between the two.[36] This research method aims to generate knowledge through the act of creation, with the creative process itself becoming a mode of inquiry. Although there are fundamentally different definitions of research-creation, the approach is defined and accepted by Canadian institutions such as the Social Sciences and Research Council of Canada (SSHRC) and the Canada Council for the Arts (CCA). For more

[34] Oxford English Dictionary, "Semantics" (Oxford University Press, July 2023), Oxford English Dictionary, https://doi.org/10.1093/OED/2159347396.

[35] John R Searle, "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–24; for a discussion of Searle's argument related to musical agents, see Eric Lewis, *Intents and Purposes: Philosophy and the Aesthetics of Improvisation* (Ann Arbor: University of Michigan Press, 2019), 57–102, https://proxy.library.mcgill.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=2046031&scope=site.

[36] Tone Pernille Østern et al., "A Performative Paradigm for Post-Qualitative Inquiry," *Qualitative Research* 23, no. 2 (April 2023): 272–89.

details on the development of research-creation Quebec and Canada, and contemporary approaches and methodologies, one may refer to Sophie Stévance and Serge Lacasse's book *Research-Creation in Music and the Arts*.[37] Greg Bruce presents a passionate account of undertaking a doctoral project utilizing research-creation at the University of Toronto and offers guidelines for graduate students interested in research-creation methodologies.[38]

The eTube project is inherently interdisciplinary, drawing on methods and theories from both academic and creative domains while encouraging collaboration between researchers, artists, and practitioners from various fields. The methodology involves an iterative research-creation process whereby the eTube, and especially the performance constraints inherent in the instrument, shapes, guides, and informs the research-creation approach to performing with MAs and adapting the *eTu{d,b}e* framework. The research objectives as part of this iterative process include assessing the musical and artistic outcomes of the improvisations from a performers' perspective, continuing to explore different artistic possibilities that emerge through interactions with MAs during performances, and partnerships with collaborators. Additional objectives include implementing eTube controller mappings for existing MAs during improvised performances and examining the artistic, collaborative, and technological advances that arise throughout this process. Through a critical reflection on my own practice and referencing academic and philosophical literature, I will be developing a heuristic for working with improvising agents throughout this document, and especially in Chapter 6. It is important to critically reflect on our creative processes, methodologies, and the knowledge produced as it is essential for articulating the contribution of the creative work to scholarly discourse. While institutional recognition of research-creation remains a challenge, our project stands out as a fitting example of the need for this methodology.

---

[37] Sophie Stévance and Serge Lacasse, *Research-Creation in Music and the Arts: Towards a Collaborative Interdiscipline*, Sempre Studies in the Psychology of Music (New York: Routledge, 2018).

[38] Greg Bruce, "Surmounting the Skepticism: Developing a Research-Creation Methodology," *Acta Academiae Artium Vilnensis*, no. 109 (August 2023): 101–23, https://doi.org/10.37522/aaav.109.2023.162.

Chapter 2: (e)Tube Context

A conduit is a channel for transporting air or fluid, a tube for protecting electric wires, and figuratively refers to the medium which transmits knowledge.[39] In the eTube project, I present a simple plastic tube, a proto-wind instrument that has been augmented and is a conduit for studying improvising musical agents.[40] A conduit for sharing my love of improvisation and computer music. A conduit for structuring an improvisation practice. A conduit for testing bi-directional interaction with MAs. A conduit for imposing performance constraints and providing clear boundaries within which to explore creatively. This research-creation project is built upon a simple and frugally designed cylindrical instrument which is my conduit for sharing a collaborative team's process of learning, adapting, and performing with existing improvising software developed by other programmers.

Polyvinyl chloride (PVC) tubes are cylindrical, their internal dimensions remain consistent. A saxophone is made of brass and is conical, the bore becomes progressively larger towards the instrument's bell. As stated by Neville Fletcher, it is the "wide conical bore and the geometry of the mouthpiece and reed that are responsible for the tone quality [of the saxophone]."[41] One might suppose that difference in sound quality would be due to the difference between the plastic tube and the brass saxophone. However, Fletcher specifies that for "wind instruments…it can be argued that the vibrating element is the enclosed air, and that the material from which the walls are made has very little influence on the sound produced," a finding which is also consistent with experiments by John Backus.[42] Considering these statements, mounting a baritone mouthpiece

[39] Conduit, *Oxford English Dictionary* (online: Oxford University Press), accessed January 15, 2023, https://www.oed.com/view/Entry/38645?rskey=SwQMV6&result=1.

[40] I borrow Greg Bruce's description of the eTube as a "proto-wind instrument" here.

[41] Neville Fletcher, "Materials for Musical Instruments," *Acoustics Australia* 27, no. 1 (1999): 7.

[42] Neville H. Fletcher and Thomas D. Rossing, *The Physics of Musical Instruments*, 2nd ed. (New York: Springer, 1998), 711; John Backus, "Effect of Wall Material on the Steady-State Tone Quality of Woodwind Instruments," *The Journal of the Acoustical Society of America* 36, no. 10 (1964): 1881–87, https://doi.org/10.1121/1.1919286.

onto a cylindrical tube will result in a different tone quality than a saxophone, but the primary factor affecting the sound should be the different bore shape, rather than the difference in material. Anecdotally, I have found the question of material to be a controversial topic among wind players, especially since one often spends significant funds on instruments and accessories. In a blog titled "The Grenadilla Myth," clarinetist and instrument builder Tom Ridenour argues that production needs have had more influence on the materials used for instrument making, rather than sound quality considerations.[43] Throughout my years playing on tubes, I have often thought about what makes an instrument's sound beautiful and what makes an instrument "professional." I play brand-name saxophones which are, in general, considered professional makes and models. They have gold lacquer coatings and engravings which also add an aesthetic quality. Some may argue that the *feeling* of an instrument may also have an important role in sound production—a higher-quality and professional feeling instrument results in a better-quality sound from the performer. However, from the literature above, and from my own experience, researchers and artists often do not agree to what extent material affects a wind instrument's tone. Throughout this document I will discuss updates to the eTube components' design and materials which strive towards a more professional instrument. The very idea of a professional eTube may seem contradictory, since when referring to musical instruments, plastic and amateur are often considered synonymous when referring to wind instruments. There are exceptions, such as Grafton's plastic saxophones, which have been played by eminent musicians such as Charlie Parker and Ornette Coleman.[44] I find that sounds from plastic tubes and saxophone mouthpiece are beautiful and compelling, and the liberty to move and spatialize the sound has influenced my approach to performance practice and improvisation. Based on limited feedback from audiences and other artists, this sentiment is shared by others. Below I will discuss why I first performed with PVC tubes, including my first performances including tubes with the Duo d'Entre-Deux, followed by my initial performances with musical agents (MAs) and saxophone. These formative experiences all led towards developing the eTube project.

[43] Tom Ridenour, "The Grenadilla Myth," no date, https://www.rclarinetproducts.com/the-grenadilla-myth.

[44] Paul Harvey, *Saxophone* (London: Kahn & Averill, 1995), 114–15.

2.1 Tubes and Duo d'Entre-Deux

Plastic tubes have been used to extend saxophones while maintaining standard and extended performance techniques by artists such as saxophonist and improviser Sam Newsome.[45] My improvised embodied performance gestures and spatialization practice with acoustic tubes was first developed with saxophonist and media artist Nick Zoulek as the Duo d'Entre-Deux in site-specific dance performances with Wild Space Dance and co-improvised works by the duo.[46] We attached saxophone necks and mouthpieces to flexible PVC tubes which allowed Zoulek and I to play the instruments while interacting with the dancers onstage. The tubes are lighter, longer, and more flexible than an alto or tenor saxophone, presenting intriguing affordances when working with movement. We used our left hand to hold the tube close to the mouthpiece while the right hand held the sounding end of the tube, free to direct the sound in space or to move with other artists or dancers. These experiences laid the foundation for my tube practice which integrates movement, spatialized performance gestures, contemporary techniques, and improvisation.[47]

PVC tubes are simple instruments and have no tone holes or keys to adjust the pitch. Yet, the acoustic tube sound is compelling and produces a low drone with audible air sound and a grainy texture. Besides the pedal tone, the sounds that I gravitate towards involve contemporary woodwind performance techniques. I often combine multiple contemporary techniques

---

45 "Sam Newsome - Home," accessed March 23, 2024, http://www.somenewmusic.com/; *Sam Newsome - Solo Soprano Saxophone Live on WFMU - July 13, 2018*, YouTube video, 52:00, 2018, https://www.youtube.com/watch?v=hKyetZ42nn4, [timestamp 08:15].

46 "Saxophonist | Composer," Nick Zoulek, accessed February 24, 2024, http://nickzoulek.com; "Home," Duo d'Entre-Deux, accessed February 24, 2024, http://www.duodentredeux.com; *Carried Away: Wild Space Dance & Duo d'Entre-Deux @ Roulette, NYC [Excerpt 1]*, YouTube video, 04:46, 2015, https://www.youtube.com/watch?v=INmeuZ3ahl0, [timestamp 02:08]; *Luminous - Tubes Excerpt - Wild Space Dance Company & Duo d'Entre-Deux*, YouTube video, 1:22, 2015, https://www.youtube.com/watch?v=G0XABGUDFfA; *Reverberant House: Presented by Duo d'Entre-Deux - YouTube*, YouTube video, 01:00, accessed February 24, 2024, https://www.youtube.com/watch?v=uKuuD6CkHbk.

47 Tommy Davis et al., "eTu{d,b}e: Case Studies in Playing with Musical Agents," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Mexico City, Mexico, 2023), 268–76, http://nime.org/proceedings/2023/nime2023_39.pdf.

simultaneously or in succession, including slap tonguing, multiphonics, singing while playing, circular breathing, flutter tongue, and air or glottal sounds. To perform these sounds, one must maintain a flexible embouchure to combine multiple techniques or switch quickly between them, described as a "'virtuosity' of the embouchure" by Marcus Weiss and Giorgio Netti.[48] Although the tube is a simple instrument, the techniques that I use are advanced techniques borrowed from my training as a contemporary classical saxophonist.

2.2 First Musical Agent Performances with Saxophones

I had been interested in improvising software since saxophonist Joshua Hyde introduced me to _derivations by Benjamin Carey in 2012. Hyde and I were living and studying in Paris at the time, and Hyde was working on _derivations with Carey, who was undertaking a PhD at the University of Technology in Sydney, Australia.[49] Hyde sent me a version of _derivations, which I was able to use with some success. When I began my doctorate in 2019, my research topics were live electronics, contemporary saxophone techniques, and improvisation and I planned to focus on live electronics repertoire including improvisation. For my second-year recital in spring 2021, I wanted a program that would showcase different approaches to live electronics and improvisation. I was interested in revisiting _derivations and my committee members suggested I investigate the OMax family of MAs and Sergio Kafejian's Construction Tools for Interactive Performance (CTIP), which led to the discovery of many MA systems.

[48] Marcus Weiss and Giorgio Netti, *The Techniques of Saxophone Playing* (Kassel: Bärenreiter, 2010), 153.

[49] Benjamin Leigh Carey, "_derivations and the Performer-Developer: Co-Evolving Digital Artefacts and Human-Machine Performance Practices" (PhD thesis, Sydney, Australia, University of Technology, 2016), https://opus.lib.uts.edu.au/handle/10453/43452.

Vincent Cusson is a programmer and digital luthier based in Montreal.[50] We met in 2019 during my first term at McGill in a seminar on timbre as a form-bearing element in music given by Professor Stephen McAdams. On October 11th, 2019, Cusson attended a Duo d'Entre-Deux concert at the White Wall Studio in Montreal hosted by *Codes d'accès* where Zoulek and I performed on saxophones and PVC tubes.[51] Cusson was intrigued by our improvisation practice and sonic palette with the tubes. Prior to my second-year recital I reached out to Cusson, a master's student in Music Technology at the time, to collaborate on my doctoral recital performance and for technological support.

2.2.1 2021 North American Saxophone Alliance Region 10 Conference Online

My first public performance with a MA was for the 2021 North American Saxophone Alliance (NASA) Region 10 Conference, hosted online by Marie-Chantal Leclair at the Schulich School of Music of McGill University (see Appendix A). Leclair and I performed with an MA called the Creative Dynamics of Improvised Interaction (DYCI2) by Jérôme Nika.[52] We performed on soprano saxophones with one DYCI2 agent trained (see Section 3.2) on a recording of me performing *Hard* (1988) for solo tenor saxophone by Christian Lauba (b. 1952). We used a DPA Microphones d:vote 4099 condenser microphone clipped to my soprano, which supplied audio input for DYCI2's audio analysis function.[53] Cusson suggested that we use a monitor screen to display the MA's amplitude output as a waveform. The screen showed a waveform that reacted in real-time to the MA's output and was larger when the MA's output was louder (see the cited

[50] Vincent Cusson, "Home," Vincent Cusson, accessed February 24, 2024, https://vincentcusson.github.io/.

[51] "Cod:A-19," Codes d'accès, accessed March 3, 2024, https://codesdacces.org/evenement/coda-19/.

[52] Jérôme Nika et al., "DYCI2 Agents: Merging the 'Free', 'Reactive', and 'Scenario-Based' Music Generation Paradigms," in *International Computer Music Conference* (Shanghai, China, 2017), https://hal.archives-ouvertes.fr/hal-01583089.

[53] "4099 Instrument Condenser Microphone," DPA, accessed March 23, 2024, https://www.dpamicrophones.com/instrument/4099-instrument-microphone.

video).[54] We had hoped that this visual aid would help the audience to differentiate between the MA's output and the acoustic soprano saxophones. Following this performance, we stopped using the screen to visualize the MAs, and only used lighting visualizations for one event, the *live*@CIRMMT concert in 2022 (see Section 4.2.4). However, we do intend to revisit lighting as one way to represent the MAs in the future. This initial performance was part of our preparation for my doctoral recital a few months later.

2.2.2 Second Doctoral Recital

My second doctoral recital program featured four different MAs and two works, which each showcased a different approach to human-computer interaction with electronics and improvisation.[55] This recital was performed on May 26, 2021, in Tanna Schulich Hall at McGill University (see program in Appendix B).[56] You will see on the program that Nick Zoulek, Kevin Gironnay, and Marilène Provencher-Leduc are listed as performers for the DYCI2 performance. Since this recital was during the COVID-19 pandemic, I decided to use recordings of collaborators from professional projects as corpus recordings for DYCI2. Provencher-Leduc and Gironnay were present in person since we also performed Gironnay's piece *Ma* to close the program. However, Zoulek is based in Chicago and was not present for this performance. This was an idea that came about because of the pandemic and the limitations of in-person collaborations. As will be discussed throughout this document, once we established the eTube project, we decided to use corpora recorded only by collaborators directly involved in the project.

---

[54] *DYCI2 by Jérôme Nika - Two Soprano Saxophone and Computer Agent Improvisation*, YouTube video, 7:18, 2021, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

[55] See McGill Doctoral Recital 2 Media: "Dissertation Media - YouTube," accessed April 30, 2024, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

[56] "Doctoral Recital: Thomas Davis, Saxophone," Schulich School of Music, accessed March 3, 2024, https://www.mcgill.ca/music/channels/event/doctoral-recital-thomas-davis-saxophone-329499.

At first, I was hesitant to program this recital. I was concerned about how to frame a program with four different improvising software, since at the time, I did not realize how differently each system would interact in performance. In addition, I was uneasy about the technological hurdles. I am thankful for the confidence of my doctoral committee who supported the program and concept for this recital. The experiences, artistic directions, research, and collaboration that was nurtured during the preparation for this performance laid the foundation for my dissertation research. The amount of preparation necessary to learn to perform with four different MAs was a considerable undertaking. Cusson spent a significant amount of time with me in rehearsals, learning about and testing the MAs, and troubleshooting technological issues. He also dedicated time to updating and testing the MAs outside rehearsals, and I am grateful for his generosity, artistry, and expertise which continues to this day. In addition, the developers of these software I performed for my recital generously supported this endeavour through online meetings, technical updates, troubleshooting, and artistic support which was integral to the recital's success.

Cusson and I worked with the MAs between December 2020 and May 2021, which occurred during the COVID-19 pandemic and lockdowns. At the time, most academic courses were held online, chamber music and in-person lessons were limited, and rehearsal space was more available than pre COVID-19. Without the dedicated time, space, and technical support to develop four MAs for this first performance, I am not certain that the project would have reached a point where I would have felt inclined to continue with it as my main research focus. This work ethic and process also set the stage for the ensuing research with the Input Devices and Musical Interaction Laboratory (IDMIL) and CIRMMT student members during the final three years of my studies. Cusson was instrumental in helping me to learn the software, updating the software, troubleshooting, and guiding and developing a shared artistic direction with me. The collaborative aspects have been very fulfilling, however the time and resources necessary to develop this project are significant and I would not have been able to commit the necessary time outside of an academic research environment. The recital preparation also developed the research-creation process and workshop routines that we would continue to expand it over the next three years as outlined in the following chapters.

Chapter 2: (e)Tube Context

The next three chapters will outline the eTube project's development over three academic years and research phases, including collaborator contributions, commissioned works, and performance examples. Chapter 3 will outline the first research phase with Cusson where we develop the beta version of the eTube controller. I will introduce and describe the existing MAs and their creators, and how we adapted them to create the *eTu{d,b}e* improvisation framework. The performance case studies will look at our first performances with the eTube. Chapter 4 examines Phase 2 and the collaboration with Kasey Pocius to create interactive spatialization models for the MAs and *eTu{d,b}e*. I will discuss the first commissioned piece for the eTube, *Enfants, apprenez-nous à parler* (2022) by Quentin Lauvray (b. 1997), and the use of MAs in this composition. I will also describe performance case studies using the spatialization software and a collaboration with the Weather Vane collective resulting in the premiere of *Improvisation Framework* (2022) by Greg Bruce and me. I will close the chapter by discussing monitoring challenges when performing a wind instrument in multichannel environments. Phase 3 is currently ongoing, and Chapter 5 will outline the hardware and firmware controller updates that Maxwell Gentili-Morin and I have completed as of March 2024. In addition, I will discuss Gentili-Morin's plans to add haptic feedback to the controller to enhance communication between the MAs and me. I will highlight Cusson's continued contributions over the past years and specific hardware updates. Finally, in the performance Case Studies I will introduce the eTube's second commissioned piece, *3tube* (2023) by Kasey Pocius (b. 1998), co-improvised by Greg Bruce, Maryse Legualt, and me, which involved adapting new eTube models for tenor saxophone and bass clarinet.

Chapter 3: Phase 1—Developing the eTube and eTu{d,b}e Framework

During the preparation for my second doctoral recital, Vincent Cusson had suggested that we research improvisation with musical agents (MA) and PVC tubes. Cusson and I were awarded a CIRMMT Student Award (2021–22) which funded our project to design the eTube controller and *eTu{d,b}e* framework. Our initial research question was: how can a simple controller facilitate interaction between an improviser and musical agents?

Cusson and I have been inspired by the many creators who have designed their own improvising systems, so rather than building our own software, we decided to adopt and adapt existing systems designed by other programmers. This chapter outlines the development of the eTube instrument, introduces the three MAs we used, and describes their implementation in the *eTu{d,b}e* framework. The development of the eTube and *eTu{d,b}e* was not a linear process, both were developed consecutively, one influencing the developments of the other, and I will attempt to outline specific details in the text below.

A word about terminology: Although I have been using the term musical agents throughout this text, at this point in the project I was not use this terminology consistently, and I only began using "musical agents" consistently in the spring of 2022 following a research exchange at the Metacreation Lab with Professor Philippe Pasquier (see Section 3.2.2). Prior to and throughout Phase 1 of the project, I would refer to the MAs with various words such as "the work," "the computer," "the piece," "the software," "DYCI2," or "agents" during rehearsals and correspondence with Cusson. As a performer new to this technological environment at the time, I was approaching MAs like any other live electronic work that I had performed in the past, with the hierarchical composer-performer and "work" assumptions carried over from experience interpreting compositions. It now seems strange that I had been referring to the agents as "pieces" or "works." As I discovered relevant literature, ongoing artistic projects using MAs, and especially following my work with Pasquier, I began to adopt language to describe the project in line with established researchers and developers in the field. This document, and much research on improvising systems, supports using language such as musical agents for describing these technologies. My reasons for doing so, including certain problematic assumptions, will be described throughout the document and especially in Chapter 6.

3.1 eTube Development

3.1.1 Infra-instrument and Constraints

Why did I choose to augment a non-typical instrument like a PVC tube, rather than a saxophone? Firstly, I was inspired by Cusson's proposition to augment the tube with a controller. We could have added a controller to the saxophone, but all digits are already necessary to engage with the instrument's keys, save the right-hand thumb. In contrast to the key-laden saxophone, the tube has no keys, leaving the hands free to engage with an electronic controller. Secondly, this was an opportunity for me to focus on a limited sonic palette, forcing myself to explore different kinds of improvised material, such as subtle timbral adjustments and long trajectories. In this regard, composer and pianist Vijay Iyer remarks that "where performers need scripts, improvisers need stimuli and constraints."[57] I had been exploring imposing constraints on my improvisations with saxophone, but the tube presents more rigid and fixed constraints. Saxophonists are known to play (too) many notes, and I am also guilty. I was interested in removing the saxophone's keywork to explicitly constrain virtuosity related to pitch. As a result, this challenged my habitual scalar, harmonic, quartertone, or multiphonic patterns I often improvised with. Instead, when performing the tube, I would rely solely on air and embouchure manipulation to adjust the pitch, overtones, or multiphonics, with much less technical "virtuosity" than the saxophone. However, this was also intriguing from a technique standpoint as it demands a certain amount of embouchure control and flexibility, as described by Weiss and Netti's "'virtuosity' of the embouchure" mentioned in Section 2.1.[58] Finally, from my experience performing with dancers alongside Zoulek as the Duo d'Entre-Deux, I was interested in the spatial and gestural

---

[57] Vijay Iyer, "Improvisation: Terms and Conditions," in *Arcana IV: Musicians on Music*, ed. John Zorn (New York: Hips Road, 2009), 172.

[58] Weiss and Netti, *Techniques of Saxophone Playing*, 153.

affordances of the tube, and what this might contribute when combined with MAs in performance.

Cusson's and my decision to augment the tube was inspired by John Bowers and Phil Archer's "infra-instrument" concept, which refers to purposefully incomplete, deconstructed, or broken instruments that are limited in terms of virtuosity and leave more latitude for electronic augmentation.[59] It was the absence of melodic- and harmonic-based notions of Eurological virtuosity which necessitated that I explore the tube's unique and rich timbral palette in lieu of my habitual performance approaches. And it is exactly this lack of pitch virtuosity which left vacant the sonic space that we decided to fill with improvising musical agents. Robert Hasegawa outlines a persuasive argument for using constraints to ignite creativity in composition, improvisation, and performance.[60] In this chapter, Hasegawa distinguishes between two types of constraints, absolute material, and relative material constraints. Absolute material constraints limit the materials one may use to a specific palette, such as specific notes in a scale, or one dynamic marking. Relative material constraints limit the relationship between musical features, such as the relationships between the harmonies that in turn govern the tonal system. Our project is primarily focused on absolute material constraints, beginning first with the tube's instrumental and performative constraints described above. These initial constraints imposed by the actual instrument will continue to inspire and guide each new direction in the project, starting with the controller design and our approach to working with MAs, and described throughout this document.

---

[59] John Bowers and Phil Archer, "Not Hyper, Not Meta, Not Cyber but Infra-Instruments," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Vancouver, Canada, 2005), 5–10, https://www.nime.org/proceedings/2005/nime2005_005.pdf.

[60] Robert Hasegawa, "Creating with Constraints," in *The Oxford Handbook of the Creative Process in Music*, ed. Nicolas Donin (Oxford University Press, online edition, 2020), https://doi.org/10.1093/oxfordhb/9780190636197.013.17.

3.1.2 eTube Controller and Mouthpiece Adapter Design

Prior to working with Cusson there were certain performance issues with the tube I wanted to address. Firstly, I would insert the mouthpiece into the tube and use transparent packing taping to secure the mouthpiece inside the tube, resulting in numerous problems. Condensation buildup would loosen the tape and the mouthpiece would no longer seal on the tube. The tape was noisy and would crinkle any time the mouthpiece moved, which was not ideal. It was also somewhat wasteful to tape the mouthpiece multiple times for rehearsals and shows. Secondly, after performing the tube for some time, my left hand, which held the tube closest to the mouthpiece, would get tired.[61] I would often move, shake, or rotate the tube with my right arm while also making sound. To secure the mouthpiece in my embouchure, I would have to steady the tube with my left hand, which was fatiguing for my arm and hand. These issues were considered during the design process described below through which Cusson found fitting solutions to these concerns.

The eTube is an augmented instrument made of a 2.54 cm diameter, 219 cm long cylindrical PVC tube augmented with a two-button controller and fitted with a baritone saxophone mouthpiece. The eTube maintains the same acoustic functions as the tube, however, the added controller is a technical and conceptual leap which obliged a proper name for this instrument. The eTube controller is designed and built by Cusson in collaboration with me and is conceived to facilitate interaction with MAs. A controller prototype may be seen in Fig. 3.1 below and the eTube beta version, which we would use for performances over the following two years, in Fig. 3.2. The buttons connect to a wireless controller which communicates with the MAs. This allows the performer another level of interaction in performance in addition to the auditory feedback from the MA's output from the loudspeakers. Cusson chose not to influence the MA's interactive settings during performances, so we could better assess the success of the controller's influence on the agents. This is not to underestimate Cusson's role during concerts: he remained a crucial

61

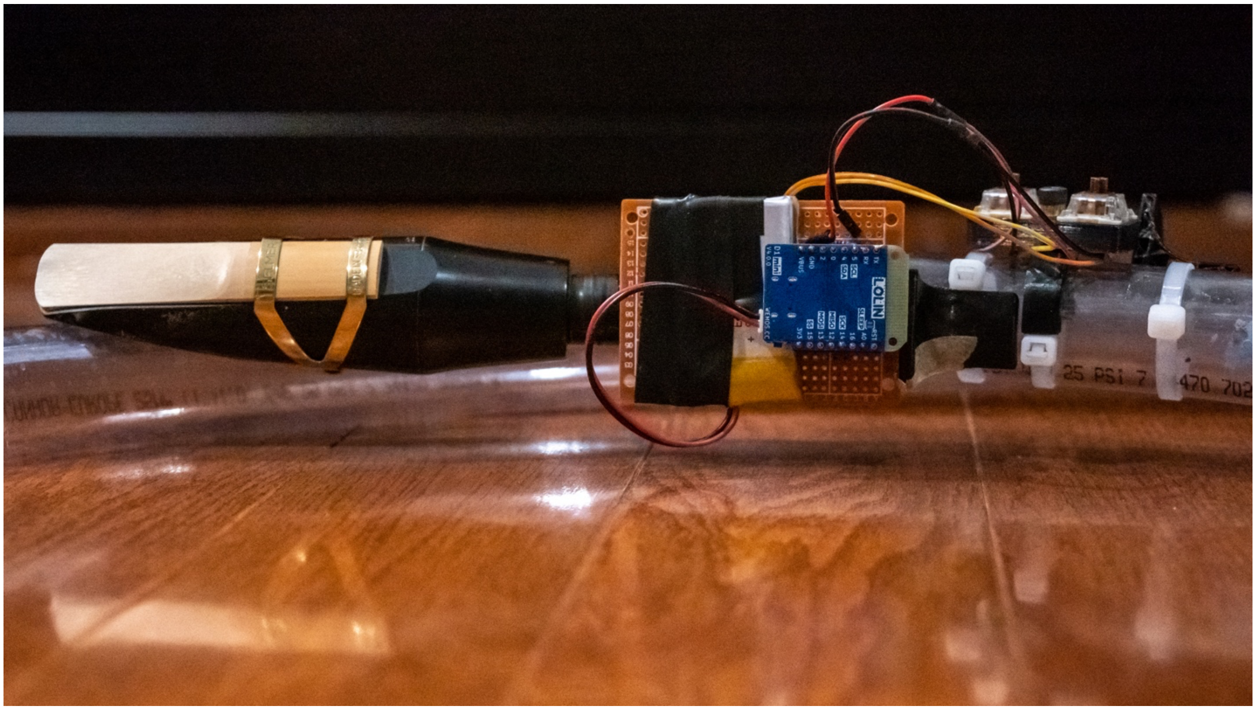

Fig. 3.1: eTube alpha version with controller and baritone saxophone mouthpiece.

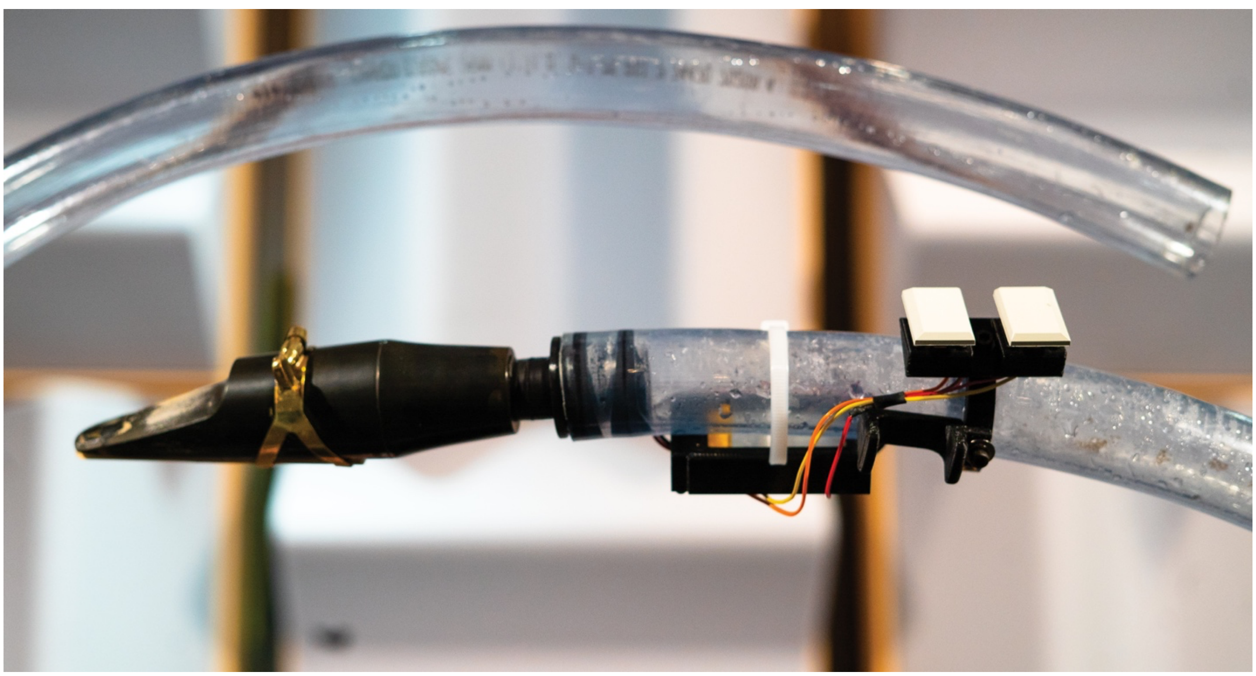

Fig. 3.2: eTube beta version with controller and baritone saxophone mouthpiece.

part of every performance as he would continue to manage and troubleshoot technological issues in real-time. We would often decide on an overall form, involving alternating between CTIP and DYCI2, or using both simultaneously, and Cusson would implement these changes in performances. In addition, he would adjust certain effects settings in CTIP, such as the playback speed of the recording module, altering the playback speed of material I recorded live.

Fig. 3.3: The author performing the eTube with Spire Muse at Simon Fraser University.

Cusson built upon my tube practice established with the Duo d'Entre-Deux to create a controller design inspired by the saxophone's keys and right-hand thumb rest. We hoped that by building the controller based on the saxophone would help a performer transfer existing instrumental technique, to minimize the learning curve. The left-hand thumb supports the instrument via the eTube's thumb rest, which allowed me to comfortably support the instrument with one hand and to secure the mouthpiece in my embouchure. The left-hand index and middle fingers are used to press buttons 1 and 2 respectively. The specific gestures and controller mappings will be discussed in Section 3.3.3. The buttons produce a clicking sound when the key cap contacts the switch below, providing both tactile and auditory feedback for the performer. The right hand typically holds the sounding end of the instrument and is free to move it in space or interact with another artist. See Fig. 3.3 above for my typical eTube playing position.

The controller uses available low-cost electronics including 3D-printed parts to secure the electronics and anchor the controller to the tube. Cusson chose computer keyboard switches for the keys since they are easily available and dependable. An ESP32 Lolin D32 board is connected to the key switches via soldered wires. A small lithium polymer (LiPo) battery is enclosed in a plastic battery holder below the ESP32 board and powers the unit. Please refer to Cusson's GitHub page to access detailed plans, parts lists, and code referenced throughout this section.[62] The 3D plans are parametric, which means that when one dimension is changed in the plan, all other dimensions are adjusted in proportion, allowing the parts to be scaled and adapted for other uses. A custom 3D printed mouthpiece adapter designed by Cusson ensures the mouthpiece is securely connected to the eTube. See Fig. 3.4 below for 3D models of the printed controller parts.

[62] Vincent Cusson, "Vincent Cusson/eTube," April 30, 2023, https://github.com/VincentCusson/eTube; See 18:14-19:50: *Ensemble AKA + Louis Beaudoin - de La Sablonnière [No Hay Banda]*, YouTube video, 25:15, 2017, https://www.youtube.com/watch?v=9rotwgh_ZTE.

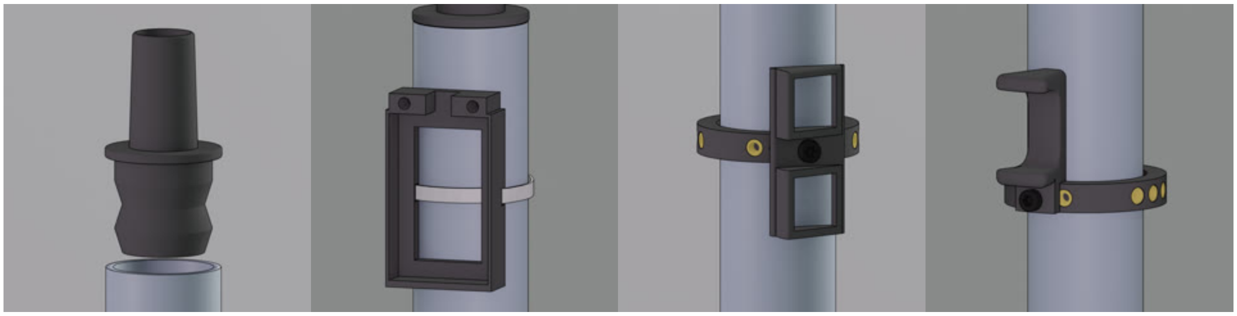

Fig. 3.4: Above are 3D models of (from left to right) the mouthpiece adapter, battery holder, switch holder, and thumb rest designed by Cusson.[63]

The mouthpiece adapter is 3D-printed from plastic (see Fig. 3.5). The top half is the mouthpiece fitting, which is conical and based on the dimensions of a baritone saxophone neck. The tube fitting is on the bottom and is designed to fit in a 2.54 cm PVC tube. Electrical tape is added to both fittings to ensure a secure fit for the mouthpiece and into the tube. As discussed in detail in Section 3.3.4, the eTube has a natural curve and to minimize torsional pressure on the hand or mouthpiece while performing, the mouthpiece must be placed in the correct orientation. The adapter is an improvement over tape as it allows the performer to easily adjust the mouthpiece angle to find the most comfortable position to hold the eTube. Early versions of the adapters broke easily, especially when I would rock the mouthpiece back and forth to remove it from the tube. The top half would break off where the mouthpiece fitting meets the tube fitting (see Fig. 3.6). Rather than printing the adapter in a vertical orientation, the adapter was printed at a 45-degree angle.[64] This resulted in the printer laying down the plastic fibre running through the problematic joint, rather than parallel with it, increasing the strength (see Fig. 3.7). A seemingly simple part, the mouthpiece adapter is a significant improvement over my use of packing tape to attach the mouthpiece. It has made performing with PVC tubes more feasible as the mouthpiece is secure, maintains a seal, and it is easier to adjust the mouthpiece angle for my embouchure.

[63] Vincent Cusson, "Vincent Cusson/eTube," April 30, 2023, https://github.com/VincentCusson/eTube.

[64] Thank you to Travis West at IDMIL who made this suggestion.

Fig. 3.5: A 3D-printed plastic mouthpiece adapter for a baritone saxophone mouthpiece.

Fig. 3.6: The mouthpiece adapters would often break in the same place.

Fig. 3.7: This baritone saxophone mouthpiece adapter was 3D-printed at a 45-degree angle.

I thought it necessary to create certain resources for composers or other collaborators to understand and work with the eTube. Below are the overtones that are accessible by overblowing on the eTube (see Fig. 3.8), from a transcription by Quentin Lauvray. Acoustically, the notes accessible by overblowing resemble the overtone series with certain overtones absent and non-typical tuning for other overtones. The eTube's fundamental note is a D1. The next usable overtone is a C4, the seventh overtone in the harmonic series, although this note is quite flat.[65] The second through sixth overtones are not accessible. D4, which would be three octaves above the fundamental, is not accessible, along with F♯4. The 11th overtone is a tritone (-49c) above the fundamental, which is not far from the eTube's G4, a fourth (+20c) above the fundamental. It is interesting to note the eTube's next overtone is a B♭ (-10c), which is similar to the 13th overtone of the harmonic series.

Fig. 3.8: The overtones accessible by overblowing a 2.54 cm diameter, 219 cm long eTube with a baritone saxophone mouthpiece, transcribed by Quentin Lauvray.

The following are some of the foundational sounds and techniques that I use when performing with the tube. Please see the *eTube Techniques Demonstration* video to listen to the core sound-producing approaches that I use with the eTube (see Fig. 3.9).[66] Movement and gesture are not explicitly included in this demonstration video, but as already stated, this is an important aspect of my practice, and may be seen in any of the videos accompanying the Case Studies.

---

[65] I specify usable overtone here because I have occasionally performed an F#3, the fifth overtone. However, I feel that it is not currently reliable enough to be included, although it may be in the future.

[66] *eTube Techniques Demonstration*, YouTube video, 0:55, 2023, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

a. Fundamental tone
b. Fundamental tone with harmonics
c. Isolated vibration modes (see Fig. 3.8)
d. Multiphonic
e. Singing while playing
f. Slap tongue articulations
g. Movement and gesture

Fig. 3.9: Some foundational eTube sounds and techniques.[67]

Included here are certain contemporary saxophone techniques, such as multiphonics, singing while playing, and slap tongue articulations. I will often perform multiple simultaneous contemporary techniques to create varying textures, layers, and beating or grain in the sound. Circular breathing is an integral part of my practice as a saxophonist and improviser and I often circular breath to extend phrases beyond what is possible with a single breath. Movement and spatialized performance gestures are also critical to my eTube performance practice and will be described in detail in Sections 3.3.4 and 6.7. As pointed out by Marcus Weiss and Georgio Netti in *The Techniques of Saxophone Playing*, it is too complex to describe how each of these contemporary techniques interact with each other, and all the resultant sound possibilities when performed together.[68] As such, I follow their guidance and list certain basic techniques, rather than attempting to describe all the possible combinations I use when improvising. In Chapter 5, I will describe the eTube's expansion to include mouthpiece adapters for bass clarinet and tenor saxophone mouthpieces, including new performance practices and techniques contributed by other performers.

---

[67] I use the word multiphonic to describe two or more pitches sounding simultaneously. A woodwind multiphonic is most often produced by specific fingerings which result in multiple pitches that are often inharmonic in nature. Since I use only my embouchure and no fingerings to produce these sounds, the technique I perform with the eTube could also be described as a split tone.

[68] Weiss and Netti, *Techniques of Saxophone Playing*, 8–12.

3.1.3 The "Ear" of the MAs: Microphone Setup

During the initial research phase, we focused on developing the controller and interaction between the improviser and MAs via controller mappings (see Section 3.3.3).[69] Following the controller development, we became more focused on researching the microphone as a sensor and how this interfaces with my performance gestures and movements with the eTube. To maintain consistency, our typical setup is one Electro-Voice RE20 microphone placed on a stand in front of the eTube performer at approximately waist height.[70] The RE20 is often used in radio or broadcast and has a heavy-duty internal pop filter, allowing it to handle the eTube's direct air pressure, and its directionality allows it to reject speaker noise without feeding back. The mic captures and transmits the improviser's live signal to the MAs, and enables the audio descriptor analysis of that signal, which influences the MA's interaction and audio output. In this sense, the microphone could be considered as the "ears" of the agent since this is how the audio signal from the eTube is sent to the agents to be analyzed. The supple nature of the eTube allows the performer a spectrum of movements, subtle to large and gentle to rapid, in addition to various orientations around and on the body, when interacting with the stationary mic. Including a stationary microphone onstage as the primary way to interact with the MAs now limits my movements to a localized area around the microphone. If I move too far away, no sound will reach the MA via the microphone, eliminating the opportunity for interaction. This has effectively adapted my practice moving with tubes through space, with dancers or other artists, into a localized performance practice with the mic as a focal point which my performance gestures gravitate towards (for additional details see Section 6.7).

---

[69] Davis et al., "Case Studies."

[70] "RE20 - Broadcast Announcer's Microphone with Variable-D by Electro-Voice," accessed April 15, 2024, https://products.electrovoice.com/na/en/re20/.

3.2 Musical Agents Introduction

This section outlines the MAs we used during the 2021–22 academic year including the *Dynamiques créatives de l'interaction improvisée* (DYCI2), CTIP, and Spire Muse. Our primary focus was on DYCI2 and CTIP during this first year, before investigating Spire Muse in spring 2022. Throughout this process, Cusson and I updated and adapted the MAs in collaboration with the developers, specific details of which will be explained below. These programs operate in Max/MSP, a graphical programming interface for realizing live electronic music.[71] Often referred to simply as Max, this environment allows one to combine pre-designed building blocks, to schedule real-time tasks, and manage communication between them.[72] MAs consist of real-time synthesis algorithms and an interface with sensor technology mapped to these algorithms allows a performer to interact with the MAs in real-time, adjusting settings which were previously only accessible on the computer.[73]

DYCI2 and Spire Muse both use pre-recorded audio corpora and are closer to the interactive end of the interactive-reactive spectrum, whereas CTIP uses effects processing and is closer to the reactive end (see Section 1.3).[74] An audio corpus is a recording that has undergone analysis, which is often referred to as a "training" phase.[75] DYCI2 and Spire Muse have different training modules and settings, but the basic process remains the same. First, the audio file is cut up into

[71] "Cycling '74," accessed April 25, 2024, https://cycling74.com/.

[72] Miller Puckette, "Max at Seventeen," *Computer Music Journal* 26, no. 4 (2002): 31, 39, http://www.jstor.org.proxy3.library.mcgill.ca/stable/3681767.

[73] Joseph Malloch and Marcelo Wanderley, "Embodied Cognition and Digital Musical Instruments: Design and Performance," in *The Routledge Companion to Embodied Music Interaction*, ed. Micheline Lesaffre, Pieter-Jan Maes, and Marc Leman, 1st ed. (New York: Routledge, 2017), 438–47, https://doi-org.proxy3.library.mcgill.ca/10.4324/9781315621364.

[74] Tatar and Pasquier, "Musical Agents," 61.

[75] Jérôme Nika et al., "Dicy2 for Max" (Ircam UMR STMS 9912, December 2022), 5, https://hal.science/hal-03892611; Kıvanç Tatar and Philippe Pasquier, "MASOM: A Musical Agent Architecture Based on Self Organizing Maps, Affective Computing, and Variable Markov Models," in *Proceedings of the 5th International Workshop on Musical Metacreation*, 2017, 1.

segments based on loudness, specifically the attacks and transients which start a note help determine the beginning of a segment. These segments are then analyzed using audio descriptors, which analyze for qualities like pitch class or timbre. In DYCI2, the user chooses descriptors from a list, and in Spire Muse, the descriptors are predetermined in the software. The audio segments are then clustered in a multi-dimensional space. Finally, segments that are closely related in the multi-dimensional space are given the same labels and organized into clusters. This analyzed data is stored with the segmented audio in a database which comprises the corpora.

In the DYCI2 documentation, an analyzed corpus is referred to as a "memory," which is a useful word for understanding the function of the corpora.[76] In performance, the agents use the same audio descriptors used to analyze the corpus, to analyze a live sound input via a microphone. The live sound is analyzed and labelled, and the agent uses this label to inform which segments from its memory will be used as an output. A real-time statistical procedure uses variable Markov models (VMM) to determine the order of the audio segments the MA will output.[77] This process allows the MAs to output "optimized responses of musical materials" since the VMM references past material and an analysis of the incoming signal to inform the MA's real-time output.[78] The MA then (re)combines the chosen segments in its corpus and outputs these as concatenated audio sequences, a process known as concatenative sound synthesis (CSS).[79]

Below, I will address specific functions and settings in each MA that have been pertinent for this discussion. Please see the cited literature for a comprehensive overview and specific details of each agent.

---

[76] Nika et al., "Dicy2 for Max," 5.

[77] G. Assayag and S. Dubnov, "Using Factor Oracles for Machine Improvisation," *Soft Computing* 8, no. 9 (2004): 604–10, https://doi.org/10.1007/s00500-004-0385-4.

[78] Nika et al., "Dicy2 for Max," 6–7, and 9.

[79] Diemo Schwarz, "Concatenative Sound Synthesis: The Early Years," *Journal of New Music Research* 35, no. 1 (March 1, 2006): 3–22, https://doi.org/10.1080/09298210600696857.

3.2.1 DYCI2

The *Dynamiques créatives de l'interaction improvisée* (DYCI2) is developed at the *Institut de recherche et coordination acoustique/musique* (IRCAM) by Jérôme Nika and colleagues and is built on OMax, SOMax, and ImproteK.[80] Interactive settings for the MAs are selected on the UI such as call-and-response, instant response, and delayed response which affect when the agent's output occurs in relation to the performer's input. Nika describes his approach to guiding DYCI2's behaviour with the phrases "follow my steps" and "follow that way."[81] These conceptual phrases were useful to understand the different interactive functions in the C–Queries module (see Fig. 3.10). "Follow my steps" is akin to "repeat last label," where the MA is outputting phrases with consistent or similar labels to the analysed input, often resulting in a more homogeneous output. "Follow that way" refers to four selections in the menu below "repeat last label" which all suggest that the MA should take the performer's current analysis as a reference point, and then diverge to varying degrees with contrasting material. How much divergence will depend on limitations of the VMM, and the corpora being used. For example, the "Last label as starting point" selection will start by outputting a segment with the same label as my live sound that was just analyzed. Following this first label, the VMM will output the most common segments which follow this original label in the analyzed corpus. In other words, the VMM keeps certain musical syntax from my original improvisation. The ways the B–Live Analyzer Input functions in performance depends on the corpora being used, and how the corpora's analysis relates to the real-time analysis of my acoustic sound.

[80] Benjamin Lévy, Georges Bloch, and Gérard Assayag, "OMaxist Dialectics," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (University of Michigan, Ann Arbor, 2012), https://hal.science/hal-00706662/; Laurent Bonnasse-Gahot, "An Update on the SOMax Project," *Ircam-STMS, Tech. Rep*, 2014; Jérôme Nika, Marc Chemillier, and Gérard Assayag, "ImproteK: Introducing Scenarios into Human-Computer Music Improvisation," *ACM Computers in Entertainment* 14, no. 2 (2017): 1–27, https://doi.org/10.1145/3022635; Nika et al., "DYCI2 Agents."

[81] Nika et al., "DYCI2 Agents."

Fig. 3.10: The DYCI2 interface with three agents including Cusson's added timer bar at the bottom.

DYCI2's UI has multiple agents which may each be trained with a separate corpus and individual interactive settings. Agent training takes place offline, and one must train each agent individually by clicking through the menu or using presets. The user must select the audio descriptors used in the analysis from a menu. This menu includes many descriptors, which can be intimidating, especially if one is unfamiliar with what the descriptors refer to, and what effect the descriptor might have on the corpus analysis. We often use the Chroma and Loudness descriptors along with other timbral descriptors such as Mel-frequency cepstral coefficients (MFCC) and Spectral Centroid (brightness). Nika comments that segmentation and analysis is "something of a dark art," and an in-depth discussion of this process, and how it might affect the interaction with MAs, is beyond the scope of this document.[82] However, those interested in more details will find useful suggestions in Nika's recent dicy2 publication (2022). I also hint at challenges and some of our solutions throughout, especially in Section 3.3.5 and the Case Studies.

In the UI, Module B–Live Input Analyzer contains an adjustable threshold level (see Fig. 3.10), which measures the incoming signal in decibels (dB). When the agent is in event mode, the threshold gate is set to a minimum input level, which launches the MA's output when the input value exceeds the threshold. There is not currently a limiter on a maximum value which would prevent an agent from launching if the input was above a certain value. Since DYCI2 has multiple agents which each may have a different corpus, the threshold function enables one to compose the corpora into specific structures or layers. For example, I would set a lower threshold for the agent I want to sound more often which might have a rhythmic or more sustained material in the corpus. I might set a higher threshold for agents that I want to sound only following my loudest gestures, such as a corpus with more accented interventions. This organization also affects the way I improvise. I will need to play more softly to only launch the agent with more sustained material and at times I might increase my dynamics gradually until the agent with accented interventions is heard. I will discuss interacting with the microphone in more detail in Section 6.7 below.

DYCI2 is defined as a library, which may be adapted or integrated into various uses such as composition or improvisation.[83] Although we have combined DYCI2 with other MAs in novel ways, we have only made minor changes to the actual Max patch, and we currently use modified versions of the tutorial patches, which are available for download.[84] In addition to other

---

[82] Nika et al., "Dicy2 for Max."

[83] Nika et al., 3.

[84] "Dicy2 | Ircam Forum," accessed March 28, 2024, https://forum.ircam.fr/projects/detail/dicy2/#project-versions-anchor; "DYCI2/Dicy2," accessed March 28, 2024, https://github.com/DYCI2/Dicy2.

programming changes, Cusson also added a timer at the bottom of the interface which indicates how long the MA's current output will be (see Fig. 3.10). The timer helped us to learn more about the UI settings. For example, if there was no sound, but the bar was moving, we knew that the MA was still functioning, there was simply silence in the corpus material. We could then try different training techniques to minimize, or edit the audio, if we wanted less space. In addition, the timer bar helped to shed light on the lengths of the MA's outputs (queries); when we adjusted the Length Query [Query Length] function, we could see how the selected number is related to the MA's output in seconds.

3.2.2 Spire Muse

Spire Muse was developed by Notto Thelle and colleagues at Simon Fraser University's Metacreation Lab as part of Thelle's PhD research at the Norwegian Academy of Music in Oslo, Norway.[85] It is a virtual musical agent that encourages musical brainstorming and acts as a jamming partner.[86] Spire Muse is built upon the Musical Agent Based on Self-Organizing Maps (MASOM) architecture and employs one larger corpus.[87] The interface allows training the agent on an audio corpus offline. The audio will be segmented and classified using a Self-Organised Map, and a temporal model of the sequence of sound object inputs is learned. In contrast to DYCI2, where the descriptors are selected by the user, in the MASOM training module for Spire Muse, there are 55 pre-selected audio descriptors used for training.[88] Once a model is created

[85] Notto J. W. Thelle, "Mixed-Initiative Music Making: Collective Agency in Interactive Music Systems" (doctoral dissertation, Oslo, Norway, Norges musikkhøgskole, 2022).

[86] Notto J. W. Thelle and Philippe Pasquier, "Spire Muse: A Virtual Musical Partner for Creative Brainstorming," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (NYU Shanghai, Shanghai, China, 2021), https://doi.org/10.21428/92fbeb44.84c0b364.

[87] Tatar and Pasquier, "MASOM."

[88] Thelle and Pasquier, "Spire Muse."

from the corpus, the folder containing the audio file and training data is easily dragged and dropped onto the interface (see Fig. 3.11) for quick start-up.

The Spire Muse interface is minimal compared with DYCI2, the four categories that analyse incoming sound are called “influences” (rhythmic, spectral, melodic, harmonic) and are adjustable manually, or via randomized global adjustments with the “Change” button. The “Go back” button sets the influences to the previous settings. The agent's global musical behavior is controlled through three interactive modes called *shadowing*, *mirroring*, and *coupling* with a fourth mode called *negotiation* which emerges from the human-agent interaction.[89] Spire Muse uses terminology more common in western musical discourse, whereas DYCI2’s terminology is more conceptual.[90] Spire Muse also uses a VMM to vary the MA’s output, and similar processes govern the MA’s output so as to maintain certain syntax from the corpora, as described for DYCI2 above in Section 3.2.1. The degree to which the VMM departs from the analyzed input from the performer depends on the interactive modes above. *Shadowing* mode will output labels that are most similar to the analyzed label from live sound, whereas *mirroring* and *coupling* will diverge to a greater extent from that analyzed label. As with DYCI2, the extent to which this diversion occurs depends on the material in the corpus, how it is analyzed, and the performer’s live sound.

[89] Thelle and Pasquier.

[90] Lewis, “Improvised Music After 1950”; Christopher Small, *Musicking: The Meanings of Performing and Listening*, Music/Culture (Hanover: University Press of New England Hanover, 1998).

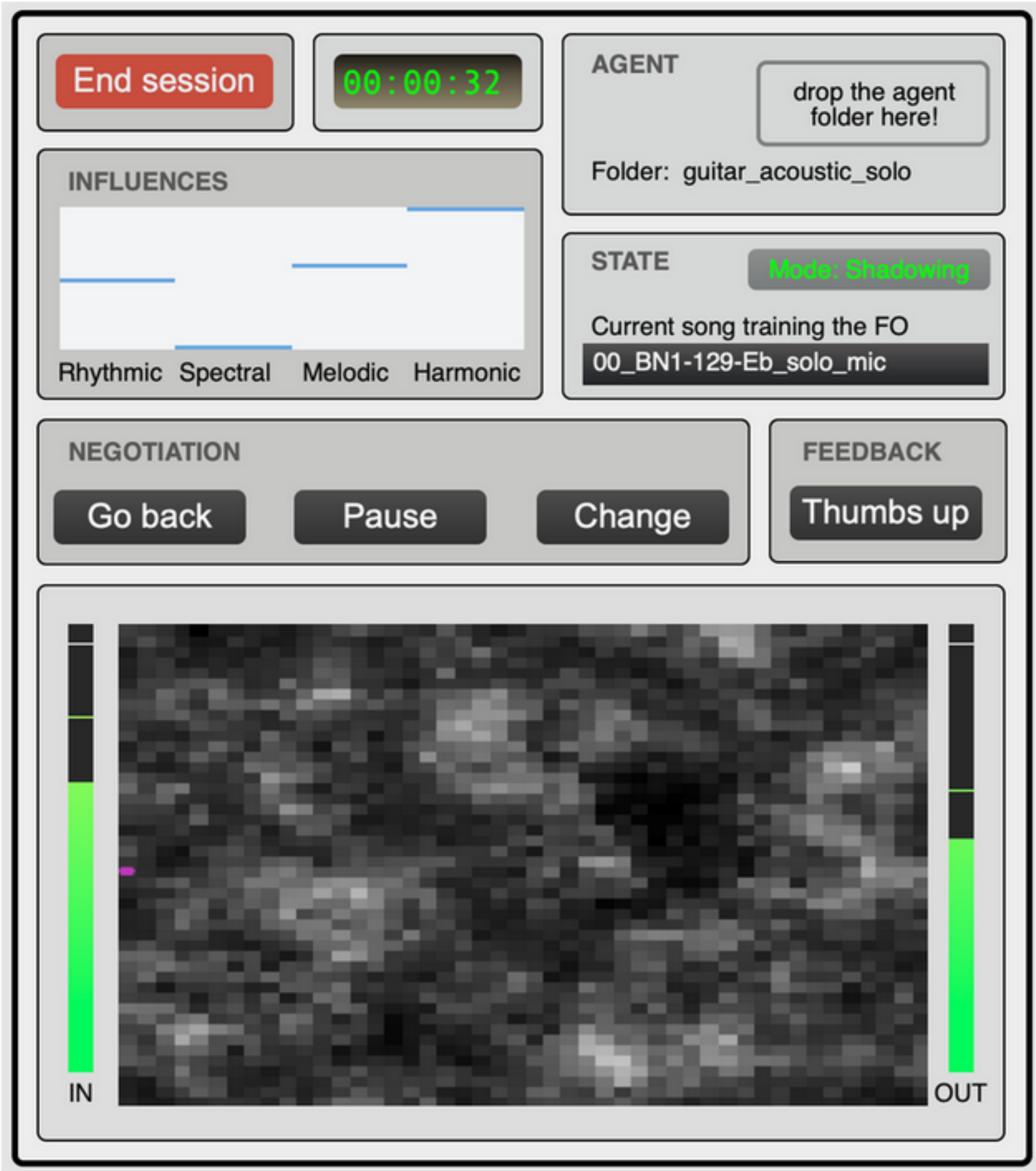

Fig. 3.11: The Spire Muse interface in negotiation mode.

In spring 2022, I was awarded a CIRMMT Inter-centre Research Exchange Award to work with Professor Philippe Pasquier at Simon Fraser University's Metacreation Lab in Vancouver, Canada. Cusson and I worked with Thelle and Pasquier to integrate Spire Muse into our framework and the exchange culminated in recorded eTube performances with Spire Muse.[91] Outcomes following this exchange include performances with Spire Muse as part of the *eTu{d,b}e* framework, and co-authoring two papers with Pasquier. The first eTube paper was published for the 2023 New Interfaces for Musical Expression (NIME) Conference held in Mexico City.[92] The second co-authored paper will be for the 2024 International Conference on Movement and Computing (MOCO) conference held in Utrecht, The Netherlands, focusing on embodied eTube gestures and interaction with the microphone.[93]

3.2.3 Construction Tools for Interactive Performance: Construction III

Construction Tools for Interactive Performance (CTIP) by Sergio Kafejian is a flexible system designed to be used for musical creation, from composition to free improvisation performance.[94] According to Kafejian, CTIP is a multi-agent modular design comprising listeners, analyzers, and sound generator modules. There are various sound generation modules including a multitrack recorder with fixed and randomized variable-speed playback, four-channel delay, granulator, and custom spectral processing effects. The listening modules analyze a performer's live sound and routes the audio to different effects modules based on a routing matrix. For

---

[91] *eTube – Spire Muse Sessions | Take 2*, YouTube video, 3:11, 2022, https://www.youtube.com/watch?v=49LuS84ZOxw; *eTube – Spire Muse Sessions | Take 5*, YouTube video, 3:57, 2022, https://www.youtube.com/watch?v=Rq5HJ07etOI; *eTube – Spire Muse Sessions | Take 6*, YouTube video, 5:41, 2022, https://www.youtube.com/watch?v=nbHG0cFNafs.

[92] Davis et al., "Case Studies."

[93] Tommy Davis et al., "Embodied eTube Gestures and Agency," in *Proceedings of the 9th International Conference on Movement and Computing* (New York, NY: Association for Computing Machinery, 2024), https://doi.org/10. 1145/3658852.3659084.

[94] Sergio Kafejian, personal communication with the author, January 15, 2023.

example, a register listener may route low sounds to the delay, and high sounds to the granulator modules. This is one way that Kafejian investigates how a system's architecture influences interaction with the system and may contribute to unity or a sense of form.[95] The output module has multichannel functions with either stereo, quadrophonic, or octophonic presets.

[95] Sergio Kafejian, personal communication with the author, January 15, 2023.

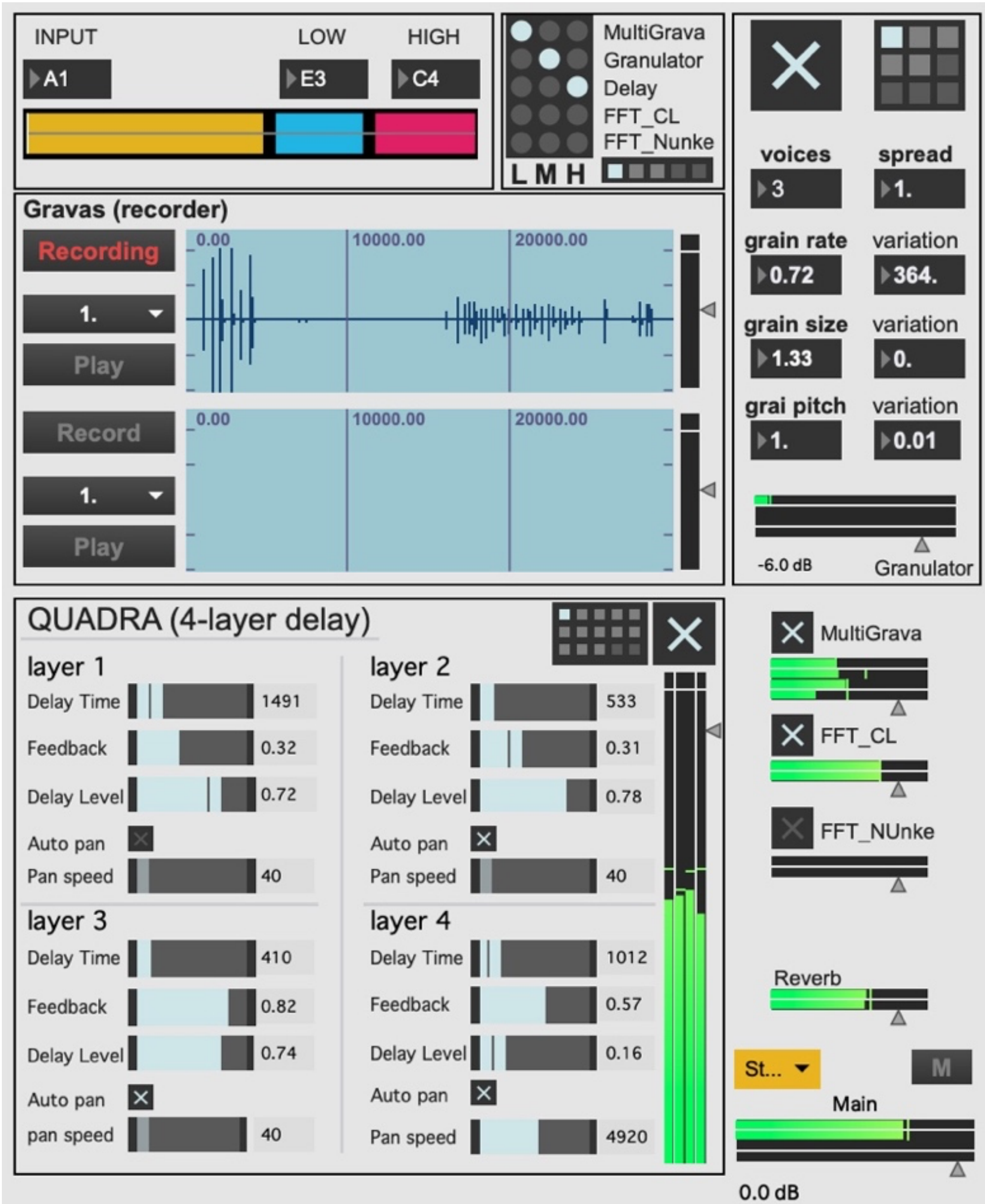

Fig. 3.12: The Construction III interface with updates by Cusson.

Cusson updated and adapted Kafejian's CTIP software, and this new version will be referred to as Construction III throughout this document (see Fig. 3.12). Kafejian designed the CTIP UI to be used by a technician or laptop performer in performances alongside an improviser. Cusson completely reorganized the UI to make it easier for me to practice with on my own, which included adding certain functions directly on the UI that were previously hidden in menus. The updated range and matrix selection tool for the pitch tracker is located at the top left of Fig. 3.11. The sliders allow me to select the three listening ranges (low, medium, high) based on concert pitch. In the matrix on the right, I can then select how the pitch tracker routes my incoming audio signal to the processing modules. In this example, low sounds below E3 are sent to the MultiGrava [Multiple Recorder], medium sounds between E3 and C4 are sent to the Granulator, and high sounds above C4 are sent to the Delay Modules. In addition, Cusson created a Main Gain Control for the stereo output, a switch to mute the main output, and reverb level adjustments accessible on the UI (see Fig. 3.12, bottom right). These small adjustments made it much easier for me to practice on my own, while having convenient access to the tools I needed, including a full mute button in case the output became too loud.

Kafejian had implemented a Multi Grava (Multi Recorder) module which automatically records and plays back live sound. For Kafejian, this module includes what Pauline Oliveros and Doug Van Nort call "episodic memory," which reintroduces material recorded from earlier in the performance. Cusson also programmed two recording modules (see Gravas [Recorder] in Fig. 3.12) which I could toggle with the eTube controller, acting similar to a loop pedal. Drop-down menus on the UI allow one to adjust the playback speed and to playback in reverse. I use the playback function to layer recorded acoustic sounds with live ones, and to contribute an element of form by playing back reproduced material from earlier in the performance. This also contributes a sense of liveness for the audience, since the module is reproducing live material, not only using fixed or pre-determined audio.[96]

[96] Simon Emmerson, *Living Electronic Music* (Burlington, VT: Ashgate, 2007), 93.

3.2.4 Algorithmic Comparison

Angèle Christin suggests "algorithmic comparison" as one approach for ethnographic studies researching algorithms as it can "shed light not only on the uses of algorithmic systems but also on their inner workings, regardless of how opaque and proprietary they are."[97] Although we did not undertake an ethnographic study comparing these algorithms, we have nonetheless gleaned specific similarities and differences through our work with them. As a result, the differences between the MAs are often inform our artistic decisions when deciding how to implement the MAs. For example, Cusson combined DYCI2 and Construction III into a single performance patch (see Fig. 3.13), and we also use Spire Muse in addition to this performance patch. With all three MAs active during performances, we may perform with all three simultaneously, pair two together, or simply use one MA at a time. This allows us to structure the improvisation and to benefit from the respective interactive and artistic approaches afforded by each MA.

I will make a general and brief comparison of what it is like to improvise with these MAs from a performer's perspective, stating some specific use cases afforded by the different MAs. David Borgo states that improvisers tend to avoid explicitly critiquing one another, however, MAs present "valuable insight[s] into the unspoken norms and cultural politics that emerge in scenes of musical improvisers" since improvisers are more likely to critique a MA's performance.[98] Borgo's perspective reinforces the importance of comparing the MAs to each other, or perhaps an imagined version of an ideal improvisation partner, but also how this process sheds light on our own artistic directions, and perhaps insights into broader questions related to improvisation between humans. Part of the process of working with MAs is considering the qualities one appreciates in an improvising partner. This project has presented an opportunity to reflect upon and learn about my own beliefs and values related to improvisation. In this sense, the MAs may act as a catalyst for considering the assumptions and cultural norms that one adopts in their improvisation and musical practice. The MAs are also imbedded with assumed cultural norms

[97] Angèle Christin, "The Ethnographer and the Algorithm: Beyond the Black Box," *Theory and Society* 49, no. 5–6 (2020): 908.

[98] Borgo, *Sync or Swarm, Revised*, 199–200.

and biases inherent in their programming and the UI, which affects the kinds of music they are more suited for.

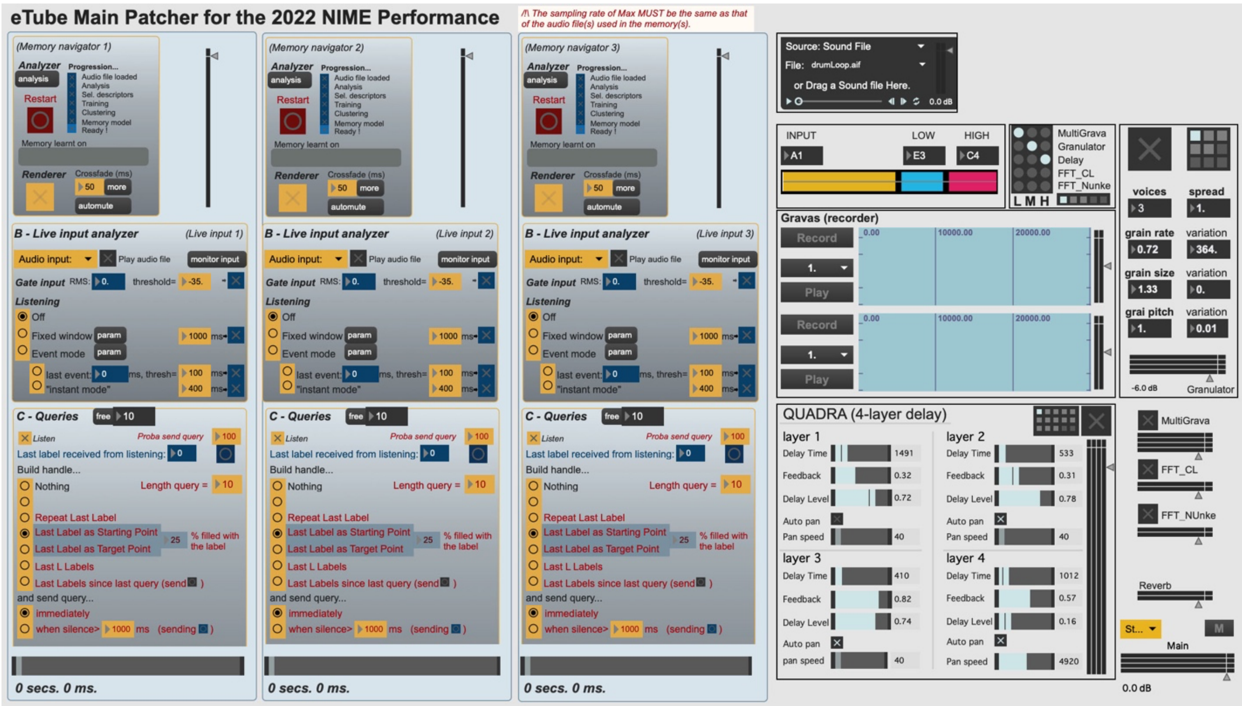

Fig. 3.13: The *eTu{d,b}e* framework performance patch including DYCI2 (left) and Construction III (right) with updates by Cusson.

As discussed by Oliver Bown and colleagues, "the claim of human-like abilities is here an obstacle to achieving the best interaction."[99] Considering the agents as equal to a human performer often leads to expectations that cannot be met by the agents, such as communication, musical development, or dynamic nuance that would be expected from a human performer (see Sections 6.1 and 6.9). Referring to specifically to MAs, George E. Lewis states that "interactions with these systems tend to reveal characteristics of the community of thought and culture that produced them." On one hand, accepting the agents for what they are—machines designed for musical interaction in non-deterministic environments and programmed to output sound in

[99] Oliver Bown et al., "The Musical Metacreation Weekend: Challenges Arising from the Live Presentation of Musically Metacreative Systems," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, vol. 9, 2013, 31.

specific ways based on analyzed sonic input—may also present new possibilities for musical interaction that are only possible with musical agents. The agents demand that we reconsider and question our own improvisation practice, including a reflection on some of the basic assumptions of improvisation, and questions that we most likely would not have encountered in performance contexts with other humans. The question of the role of bodies in improvisation, as suggested above in Section 1.2, is a critical consideration when working with MAs, and one I will address in Chapters 4 and 6. Other questions have arisen such as the nature of phrases in improvisation and what constitutes a musical phrase considering the segmentation and concatenative synthesis processes the MAs compute?

As suggested above, each MA has certain settings in the user interface, which to us suggest specific artistic affordances. For instance, an important distinction between the DYCI2 and Spire Muse corpora is the corpora length that we have decided to use. The DYCI2 tutorial interface presents three agents and we have decided to use shorter corpora for each agent. Whereas Spire Muse loads one corpus at a time, and we have decided to use one larger corpus of approximately 45 minutes with it. However, this is only how we have decided to use the MAs, one could also use a larger corpus with DYCI2, or a smaller corpus with Spire Muse, for example. In addition, the qualifiers larger and smaller are used here to denote the relative size of the corpora in the context of our project. Compared to the datasets used in large language models for AI research, both the DYCI2 and Spire Muse corpora are relatively miniscule. Since we use multiple agents with DYCI2, a performance might be compared to a group improvisation where each member has their own performance style. Therefore, with DYCI2 we often choose a different corpus for each of the three agents as a way to layer or to create contrasting responses. When performing with Spire Muse's larger corpus, it is more like a single performer who has diverse performance material. Construction III's Gravas (Recorder) function allows for the playback of audio recorded during the performance, adding an element of liveness to the performance since live sound is reproduced. Construction III's pitch tracker allows one to influence what effect will be applied based on the live sound's tessitura. For example, I might route low tones to the FFT_Nunke object, which will create a larger and more sustained texture. I would route the upper register to the Quadra (4-Layer Delay) module, which will spatialize a delay effect. Using these two effects, I may begin with a low note, creating a sonic foundation, followed by repeated harmonic squeaks in the upper register, which will then be spatialized by the Quadra Delay to create a rhythmic textural layer in the upper register. In contrast to DYCI2 and Spire Muse, Construction III may be thought of as an electronic musician who operates a variety of live processing modules. As with my approach to outlining contemporary techniques, I have only stated some basic distinctions here, rather than attempt an exhaustive list of all combinations. However, the case studies below will shed additional light on how the eTube team has combined the agents, with media examples to illustrate these choices.

3.3 eTu{d,b}e Framework

I am indebted to Vincent Cusson who named the eTube instrument and *eTu{d,b}e* framework, including the combined notions of etude and learning which have provided structure and artistic inspiration throughout this project.

3.3.1 Etude and Naming Scheme

eTu{d,b}e is the improvisation framework first developed by Cusson and me in Phase 1 of the project, and simultaneously refers to the name of the eTube and to a series of improvised etudes based on human-computer musical interactions (see Fig. 3.14). This framework adapts the MAs described in Section 3.2 in a flexible performance architecture. The curly brackets in the title are borrowed from computer terminology. This syntax is used to indicate that both letters are interchangeable, representing the intertwined relations between the eTube instrument and how we navigate the notion of etude. These details will be outlined in the case study sections and accompanied by media examples throughout the document. The curly brackets also contain the acronym for decibels (dB), a unit for measuring the relative loudness of sound. The French word *étudier* (to study) suggests how the human and agent learn from each other through performance. We intended to create different etude structures to investigate interaction with the agents using various corpora, mapping, and interactive settings. While working with the agents, the humans study the MA(s) and learn how they react differently with specific corpora or listening settings,

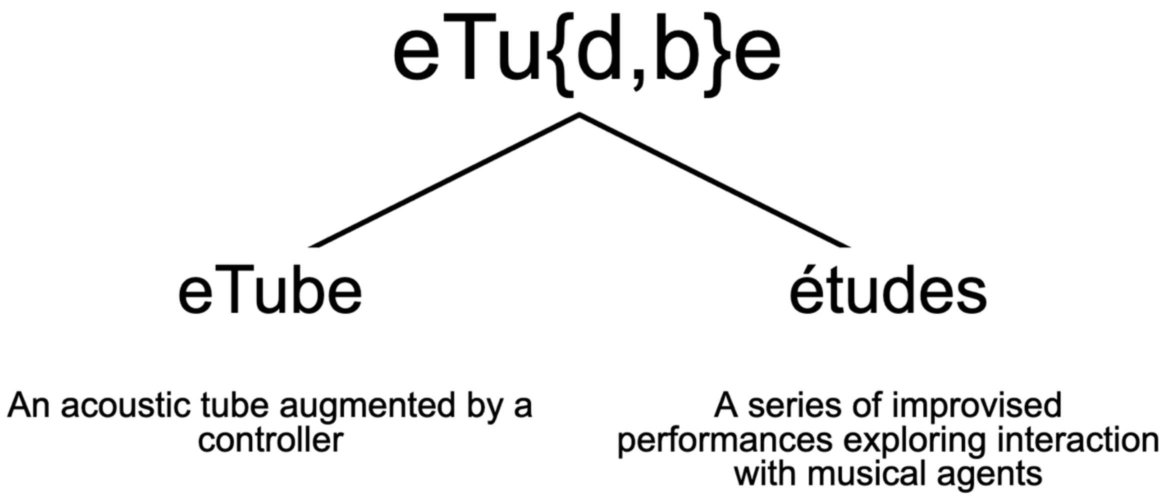

Fig. 3.14: The *eTu{d,b}e* naming scheme.[100]

and the MAs analyze the improviser's input via a microphone. This learning process may include developing a type of ear training called "algorithmic listening" where the improviser and/or programmer gains insight into the procedure causing a musical event.[101] In other words, one might develop the ability to hear the difference between various listening or interactive settings in the software solely based on the agent's behaviour and sonic output. The programmers and improvisers are always learning about the musical agent through testing and performance;[102] however, the agents we use currently do not have short- or long-term learning functions. A contrasting example is Benjamin Carey's *_derivations*, an MA that may learn both during a performance and between performances.[103] This is a desirable function in MAs since it allows

[100] Thank you to Vincent Cusson for this graphic.

[101] Andrew R. Brown et al., "Interacting with Musebots," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Blacksburg, VA, 2018), 22.

[102] Robert Rowe, *Machine Musicianship* (Cambridge, MA: MIT Press, 2001), 7.

[103] Benjamin Carey, "Designing for Cumulative Interactivity: The _derivations System," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (University of Michigan, Ann Arbor, MI, 2012), https://doi.org/10.5281/zenodo.1178227.

the agents to "evolve their behavior in relation to accumulated patterns of input."[104] While they do not learn in the course of a performance, the MAs we use analyze the audio recordings used for their output in a separate training phase. I will discuss this training phase shortly; however, it is important to note that once this training phase is complete, the analysis remains static. The training phases, especially in DYCI2, are relatively quick, and so it is possible to have the same corpora trained on different audio descriptors and segment sizes, which will affect how the corpora is output by the MA, and to manually upload these at different times during a performance.

3.3.2 Corpora Creation and Curation Process

During phase 1, corpora curation was limited to eTube and saxophone recordings performed by me, and we primarily used recordings of improvisations. We also used recordings of me performing two contemporary compositions as corpora. *Hard* (1988) for solo tenor saxophone by Christian Lauba (b. 1952) was used for the NASA performance (see Section 2.2.1), and *Le fusain fuit la gomme* (1999) for solo baritone saxophone by Marie-Hélène Fournier (b. 1963) for the McGill Saxophone Studio Recital (see Section 3.4.2). We could have continued to use my own recordings of saxophone repertoire or improvisations for the corpora, but we wished to use *eTu{d,b}e* and the corpora as a way to document my ongoing exploration of the eTube's sonic identity. The corpora recordings document certain elements of my eTube performance practice at a given time, and updated corpora also suggest what kinds of musical interactions we were interested in pursuing with the agents. We edited certain aspects of the audio files as part of the corpora curation process. However, we decided to leave in certain artefacts such as my inhalations or other bodily sounds associated with holding the eTube or preparing it to be played. When I played and sang, we could not remove my voice from the audio, and slap tongue articulations would inevitably result in bodily resonances from me and the eTube, resulting in certain corporeal shadows remaining in the corpora. As a result of this curation process, I was intimately familiar with the improvised material in the MA's corpora because I had recorded it.

[104] Garth Paine, "Interactivity, Where to from Here?," *Organised Sound* 7, no. 3 (2002): 298.

Recording and testing our corpora has been a continual evaluation process and we have integrated new artistic ideas (e.g., rhythmic motives) following performances and listening sessions. The corpora have served both as a mechanism for refining our corpora creation process and as a repository to document the eTube's sonic development as I have advanced its performance practice.

3.3.3 Controller Mapping and Design Process

We originally intended to map a global set of adjustments in the MAs we use, meaning that one controller command would adjust a host of interactive settings, significantly adjusting the MA's behaviour. However, our mappings are currently limited to one-to-one commands, such as launching an MA's output or turning on or off listening settings (see Figs. 3.15 and 3.16). Broadly speaking, we have maintained the same mappings throughout the project thus far. These mappings work well musically, I have practiced with them, and we became been busy with performances, other collaborations, and papers, and have not yet revisited our ideas for more global adjustments. There is always a balance between introducing new programming, hardware, or corpora and preparing for upcoming concerts with an instrument that functions well and allowing a performer time to practice with that instrument and system.[105] Over the next years, we would focus on developing corpora, spatialization software, and commissioned works, rather than controller mappings.

We asked the question: what would be a convincing way to influence/interact with a specific MA in improvisation? The resultant mappings came from our experience performing and working with the MAs, and the mappings are based on certain interactive features that are inherent in the software, and features that we felt would result in meaningful and musical interactions during improvisations.

[105] Pierre Alexandre Tremblay, Nicolas Boucher, and Sylvain Pohu, "Real-Time Processing on the Road: A Guided Tour of [Iks]'s Abstr/Cncr Setup," in *International Computer Music Conference*, 2007, http://eprints.hud.ac.uk/id/eprint/999/.

Both controller buttons are programmed with three different gestures: one click, a double click, and a long click. Combining these three gestures allows us 27 different inputs from the controller that may be programmed to adjust settings in the MAs. Some of the 27 combinations would be unrealistic to use (e.g., one long click of button 1 followed by a double click of button 2). However, we have found that the ones presented below have been sufficient for our needs. Rather, we embrace this limitation and planned to change the mappings for different performances depending on the corpora, collaborators, performance context, and the MA(s) used.[106]

Controller Action	**Intended Interaction**
Single click button 1	Launch output for agent 1
Double click button 2	Turn on/off agent 2's listening function
Long click both buttons	Start/stop loop recorder
Double click both buttons	Start/stop loop playback

Fig. 3.15: DYCI2 and Construction III mapping table.[107]

Controller Action	**Intended Interaction**
Double click button 1	Toggle "Change" function
Double click button 2	Toggle "Go back" function
Double click both buttons	Toggle "Pause" function

Fig. 3.16: Spire Muse mapping table.[108]

[106] Bown et al., "The Musical Metacreation Weekend," 32.

[107] Davis et al., "Case Studies."

[108] Davis et al.

3.3.4 eTube Embodied Performance Practice

My interest in working with the eTube comes from the ubiquitousness of both plastic material and noise in our lives. PVC tubing is ubiquitous in our lives and is used to transfer fluids or wastewater in domestic and industrial settings. Often hidden away in walls, ceilings, cabinets, or crawl spaces, PVC transfers detritus fluid from our homes, workspaces, and industries to other environments. Air conditioners remove humidity from the air and collect excess water before it is evacuated through PVC tubing by a mechanical pump or directly via gravity. Mechanical noise signals the extraction process as water bubbles and gurgles through the cylindrical passage. The PVC is a conduit which guides the detritus away from our environment, it also provides structure and is the resonating body that limits and directs the water's pathway, resulting in the gurgling sound from the air and water mixture as the pump drives the water out of our environment. Noise and plastic, often hidden away or out of our conscious awareness, only to be noticed when it breaks down (a leaking pipe) or becomes the focus of attention (a car alarm at night). In this project, I am interested in exploring the artistic affordances when a flexible plastic tube and noisy contemporary techniques are prioritized as principal materials in improvisation.

Developing a new augmented musical instrument (AMI) also involves developing the performance practice surrounding that instrument.[109] The PVC tubing purchased for this project was stored on large spools which resulted in the material having a natural curve that has influenced how I hold and move with the instrument. My training as a musician, and not as a mover, is apparent since I use primarily use my hands to interact with the eTube and I adopt a playing position similar to that of the saxophone (see Fig. 3.2 above).[110] However, movement of my body and the eTube have proven to be integral to the instrument's performance practice, and

---

[109] Eduardo Reck Miranda and Marcelo M. Wanderley, *New Digital Musical Instruments: Control and Interaction beyond the Keyboard* (Middleton, Wisconsin: A-R Editions, Inc., 2006), 21–25, http://site.ebrary.com/id/10857388; Jeff Kaiser, "Improvising Technology, Constructing Virtuosity," *Cuadernos de Música, Artes Visuales y Artes Escénicas* 13, no. 2 (2018): 87–96.

[110] Doga Cavdir and Sofia Dahl, "Performers' Use of Space and Body in Movement Interaction with A Movement-Based Digital Musical Instrument," in *Proceedings of the 8th International Conference on Movement and Computing* (New York, NY: Association for Computing Machinery, 2022), https://doi.org/10.1145/3537972.3537976.

the eTube's flexible nature presents intriguing affordances in this respect.[111] Composer and violinist Malcom Goldstein discusses finding a uniqueness in sound as “a kind of focusing on, making them focus on what is necessary to do, to make a sound; that is, upon their own physicality, upon the way they move, to become aware of the gesture of their sounding.”[112] As a saxophonist, removing the keys from an instrument has put my attention elsewhere, with a focus on moving the instrument in space and the sonic results from this interaction. Thus, I am much more aware of how the eTube’s orientation in space and my body’s movement with the instrument affect the acoustic sound, and also how that sound is picked up by the microphone. From my typical playing position (see Fig. 3.3), I often rotate my torso back and forth, moving the eTube in a semi-circle around my body (see cited video at timestamp 2:35–3:20).[113] In addition, I frequently spin the eTube's sounding end with my right hand in various orientations around my body including above my head and in front of the microphone. As suggested by Goldstein, these constraints have forced a certain focus, perhaps inwards towards sound, but also towards my own corporeal interactions with the instrument, and the resultant acoustic and tactile feedback I receive. In this regard, Derek Bailey suggests that “the accidental can be exploited through the amount of control exercised over the instrument, from complete—producing exactly what the player dictates—to none at all—letting the instrument have its say.”[114] I cannot perform the eTube with perfect accuracy, and I would not want to, as I find that the interactions are much more fruitful when inspiration is taken from the instrument’s propositions. When I overblow isolated harmonics, multiple notes may sound rather than the single note I was aiming for, adding a certain colour, timbre, and texture to the phrases. This may inspire me to take the

---

[111] Balandino Di Donato, Christopher Dewey, and Tychonas Michailidis, “Human-Sound Interaction: Towards a Human-Centred Sonic Interaction Design Approach,” in *Proceedings of the 7th International Conference on Movement and Computing*, 2020, 1–4, https://doi.org/10.1145/3401956.3404233; Paul Dourish, *Where the Action Is: The Foundations of Embodied Interaction* (Cambridge, MA: MIT press, 2001), 55–98.

[112] Malcolm Goldstein, *Sounding the Full Circle: Concerning Improvisation and Other Related Matters* (Sheffield, VT: M. Goldstein, 1988), 84, http://www.frogpeak.org/unbound/index.html.

[113] *eTu{d,b}e by Kasey Pocius and Tommy Davis – IIICON 2023*, YouTube video, 9:56, 2023, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

[114] Bailey, *Improvisation*, 100.

improvisation in a new direction, or to stay and explore this unanticipated sound for longer. What might be considered a "mistake" in other contexts, is taken as inspiration for musicking. Perhaps it is also partially due to the limitations of the eTube, that such unintended departures might be viewed as productive, and as welcomed opportunities to engage with other sounds.

Since the corpora are primarily eTube recordings, they are similar to the acoustic eTube, and this may blur the distinction between electronically and acoustically produced sounds for the audience. This situation may also blur the distinction between considering the eTube as an instrument augmented with electronics, versus the eTube performer and the MAs considered as two separate realities. The former refers to Robert Rowe's "instrument" paradigm whereas the latter refers to his "player" paradigm, as introduced in Section 1.3.[115] The eTube is a new instrument, and the audience may be unfamiliar with its sound palette and performance practice. Although the basic sound production is clearly like many woodwind instruments, like the saxophone, audiences would not have the same cultural reference and understanding as with other well-known instruments. The performance practice is not so far from the saxophone though, so they would most likely have some understandings based on standard woodwind performance practice. This may result in the blur between acoustic and amplified sound being magnified as the audiences may be first learning about the eTube's performance practice as they are listening to the performance. Alexander Harker and Pierre Alexandre Tremblay state that "any loudspeaker and room in combination will have an effect on the sound heard by a listener," demonstrating that amplified electric signals have certain characteristics that distinguish them from the same sounds when produced acoustically.[116] In other words, an amplified saxophone will sound like an amplified saxophone because of the specific loudspeaker and room acoustic combination, which will be different than the saxophone's acoustic sound in the same space. A fascinating experiment might consider the perceptual differences between audiences who understand the performance within Rowe's instrument paradigm versus the player paradigm, and

[115] Rowe, *Interactive Music Systems*, 8.

[116] Alexander Harker and Pierre Alexandre Tremblay, "The HISSTools Impulse Response Toolbox: Convolution for the Masses," in Proceedings of the International Computer Music Conference (The International Computer Music Association, 2012), 148–55, http://eprints.hud.ac.uk/id/eprint/14897/.

how the loudspeaker setup and amplification affect this understanding in improvisation with MAs.

3.3.5 Performance, Testing, and Artistic Evaluation

Through an iterative process of improvisations, listening sessions of past performances, and hardware design the team is continually learning about the MAs through testing and performance. Performances are both one of our research outputs, and part of our methodology for exploring interaction with the MAs and the eTube. Performances act signposts during the development process where we fix the corpora, controller mappings, and interactive features in the software, which allow a point of reference for certain variables during the process. Regarding evaluation of MAs, Oliver Bown comments that "there are no simple, objective measurables that indicate when computer generation of output has been creatively successful."[117] In this project, listening to recorded rehearsals and public performances plays an important role in the reflective and artistic evaluation processes. I specify *artistic* evaluation here because, as suggested by Bown, we are not currently evaluating the agents objectively or empirically, but rather we are evaluating them subjectively as one would in most artistic practices. These subjective evaluations are considered along with relevant research, and this helps develop specific solutions for interacting with MAs, which in turn inform future developments.

Within musical metacreation (MuMe) collaborations, Oliver Bown and colleagues differentiate between two types of agency that guide the development process, which they call "performative" and "memetic" agency.[118] Performative agency refers to the influence an MA exerts over an improviser within a concert performance and memetic agency is related to longer-term influences that a MA system might have on musical styles more generally during an offline

[117] Oliver Bown, "Player Responses to a Live Algorithm: Conceptualising Computational Creativity without Recourse to Human Comparisons?," in *The Sixth International Conference on Computational Creativity* (Park City, UT, 2015), 126.

[118] Oliver Bown, Alice Eldridge, and Jon McCormack, "Understanding Interaction in Contemporary Digital Music: From Instruments to Behavioural Objects," *Organised Sound* 14, no. 2 (August 2009): 194–95, https://doi.org/10.1017/S1355771809000296.

composition process. In the eTube project, an interdisciplinary team shares in contributing to this iterative process and new developments. Although certain contributions may be attributed specifically to one person, many are a result of our specific interdisciplinary working process which stems from everyone's specific experience and interests. Therefore, it is often difficult to attribute specific developments to one person as they often emerge through the often messy and intertwined relationships of the creative and scholarly in research-creation.

One of the challenges of working with MAs is attempting to understand and evaluate how the audio descriptors influence the MA's behaviour. Machine listening is the metaphor often used to describe the processes that model human listening, which often involve "converting sound into data and then subjecting those data to a machine learning process."[119] Researchers do not fully understand human listening, and there is no singular research field or agreement on how machine listening operates. As Jonathan Sterne outlines, the approaches are often specific to the application.[120] When choosing audio descriptors to use for the DYCI2 corpora analysis, we often choose chroma, which analyses the audio signal for pitch class. This might seem like an odd choice considering the focus I have put on noisy techniques. However, in our attempts to understand these processes over the past years, we have found that in general, the MAs return a plausible response to interactions when this this descriptor is privileged, although we may not be able to objectively detail the reasons why. Although we also do not fully understand the processes that govern these audio descriptor analyses, we have nonetheless developed our own understanding and preference for working with specific audio descriptors through experimentation. Probing the corpora models has been an interesting mode of inquiry. One might assume that, using these listening models, and by having the corpora sounds the same as those produced by the live instrument input, that the MAs should be able to find similar matches in the corpora. For example, if I perform a slap tongue on the eTube, one might assume that the MA would output segments that also include some slap tongues, since the corpora contains eTube slap tongues recorded by me. But of course, one would not want an exact replication of the live

[119] Jonathan Sterne, "Is Machine Listening Listening?," *Communication +1* 9, no. 1 (October 2022): 1, https://doi.org/10.7275/zeqh-eg38.

[120] Sterne, 1.

improvised material to avoid the MAs simply acting like a sort of looper pedal or mimicking an improviser's input too closely.

The MAs algorithms could be described as a black box, where one may observe the inputs and outputs, but the internal processes are opaque.[121] According to David Borgo, engaging with a black box requires a "fundamentally performative engagement."[122] 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. Each of the collaborators has a different perspective on the size of the black box. For me as an improviser, my conception of the MA's black box is much larger than that of Cusson's, for example. This concept will be discussed further in Chapter 6 below.

3.4 Case Studies

3.4.1 *live*@CIRMMT 2022 (Cancelled Due to COVID-19)

The first public concert with the eTube, newly designed controller, and *eTu{d,b}e* framework was scheduled for early February 2022. This concert was cancelled due to COVID-19. However, it served an important milestone for Cusson and I to work towards. This entire process of preparing a new AMI for performance was new to me, and it challenged my concert preparation process for performing repertoire on the saxophone or improvising. The disappointment experienced because of the cancelled concert provided an opportunity to pause and consider the process of preparing for this performance. Although this was an improvised performance, I was treating the whole setup as I would treat a piece to be performed. In other words, I expected to practice and play with the exact performance setup leading up to the concert. I did not

[121] William Ross Ashby, *An Introduction to Cybernetics* (London: Chapman & Hall Ltd., 1956), 86–117.

[122] David Borgo, "Openness from Closure: The Puzzle of Interagency in Improvised Music and a Neocybernetic Solution," in *Negotiated Moments: Improvisation, Sound, and Subjectivity*, ed. Gillian H. Siddall and Ellen Waterman, Improvisation, Community, and Social Practice (Durham: Duke University Press, 2016), 113–30, https://doi.org/10.1515/9780822374497.

understand the layers of work that went into developing an AMI, and the different types of practice and testing necessary throughout this process. I was used to practicing a piece the way it would be performed, and although I considered myself a flexible performer-improviser, this process of developing the eTube was challenging. I was often concerned with being able to practice with the eTube in the same setup as I would have during the recital. Everything came together prior to the cancelled concert, and I was indeed able to practice like this, but it was a long road of testing, trying, tweaking, and troubleshooting prior to arriving to a place where I could practice in the way that I was accustomed to as a performer. The question of what practicing means in the context of MAs and improvising software, or in developing interactive electronics works for that matter, was a point of reflection at the end of this academic year. I do not necessarily have concrete answers to the question of how to practice with MAs, but as I became more comfortable and engaged in this process, I came to understand the necessary steps as an ongoing practice, and the different elements of rehearsal, workshop, testing, tweaking, troubleshooting that are indeed an integral part of the practice.

3.4.2 McGill Saxophone Studio Recital Spring 2022

The first public performance of *eTu{d,b}e* within an academic setting was for the McGill Saxophone Studio recital in Clara Lichtenstein Hall on April 5, 2022.[123] This performance used only DYCI2 and Construction III, since we had not yet implemented Spire Muse. The three corpora used for DYCI2 included "all tubes," which was a corpus compiled of multiple improvisations based on specific gestures. For example, I had performed long tones followed by a slap tongue, playing and singing, articulated fundamental, and articulated harmonics. Once I had compiled these gestures into one track by hand, I pitch shifted all gestures up and down by a tritone, an octave, and an augmented 11th intervals for additional variation. I also used the "water tubes" corpus, which contains recordings of the eTube filled partially with water, creating fluctuating dynamics, unintended upper harmonic squeaks, and gurgling water noise. Filling the

[123] *eTu{d,b}e by Vincent Cusson and Tommy Davis*, YouTube video, 8:18, 2023, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

eTube with water was inspired by the materiality of the PVC as described above in Section 3.3.4. The final corpus was a recording of me performing *Le fusain fuit la gomme* by Marie-Hélène Fournier. I chose excerpts of specific sections from this piece that featured repeated articulations with timbral alterations and repeated slap tongue multiphonics. Construction III's Gravas (Recorder) module playback was set at a quarter speed, resulting in my recorded material being played back more like a rhythmic texture than as a pitch. We used a Barcus Berry contact mic to feed into DYCI2 and the Electro Voice RE20 was used for Construction III effects. As a result, the agents had a consistent feed of my acoustic sound via the contact mic, and the gestures through space and in front of the RE20 affected only the acoustic sound and Construction III's effects. I will discuss more details related to interaction with the microphone and different mic setups for concerts in the Case Studies below, and especially in Sections 4.3.4, 5.3.1, and 6.7.

3.4.3 New Interfaces for Musical Expression 2022 Conference

Vincent and I performed for the 2022 New Interfaces for Musical Expression (NIME) Conference, which was held at The University of Auckland, New Zealand from 28 June–1 July 2022.[124] I consider this the premiere performance for the eTube and *eTu{d,b}e* framework since it is the first presentation outside my academic courses at McGill. We submitted a pre-recorded improvisation with MAs recorded at CIRMMT between April 12–15, 2022.[125] We used Cusson's performance patch (see Fig. 3.13) with three DYCI2 agents and Construction III for this improvisation. DYCI2's agent one used an eTube slap-tongue corpus, agent two contained baritone and tenor saxophone slap-tongue articulations, and agent three was a mixture of eTube, tenor saxophone, and baritone saxophone slap tonguing. Tapping button two twice would turn agent two on or off. A single tap of button one toggled a musical statement from agent one. We used different threshold values in the DYCI2 listening module for each agent so that I would

---

[124] Vincent Cusson and Tommy Davis, "Etu{d,b}e: A Preliminary Conduit," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (The University of Auckland, New Zealand, 2022), https://doi.org/10.21428/92fbeb44.c05957ee.

[125] *eTu{d,b}e by Vincent Cusson and Tommy Davis - NIME 2022*, YouTube video, 14:12, 2022, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

activate certain agents with a lower volume input and other agents required a higher input. By varying my dynamics and proximity to the microphone, I could influence how many agents were triggered by each gesture. If I played softly, only the agent with the lowest threshold would output sound, and if I played strongly, all three agents would output sound. These threshold settings help to create varied responses from the three DYCI2 agents.

For Construction III the controller was programmed to manipulate a stand-alone recording module with variable-speed playback, enabling the reproduction of the performer's live sounds in concert (see Fig. 3.15). Simultaneously holding both buttons started the recording, and releasing both buttons stopped the recording. A double click of both buttons would launch the recording playback. This module had not been routed to DYCI2 to interact with the agents and remained primarily an expressive tool for the performer to reinforce, interact with, or to suggest formal structure by using playback of past improvised material

3.5 Chapter 3 Conclusion

Phase 1 of the eTube project involved first developing the eTube instrument parts and controller alongside Vincent Cusson. This instrument was influenced by the notion of the infra-instrument, which presents significant performance constraints, challenging me as an improviser to explore a more limited sonic palette than that of the saxophone. I describe our process of learning to adopt and adapt existing MAs, rather than creating our own software. We developed the *eTu{d,b}e* improvisation framework which includes three MAs: 1. DYCI2, 2. Construction III, 3. Spire Muse. The concept behind this framework is based on the idea of the etude and exploring various etudes on human-computer interaction through improvised performance with the MAs. This exploration of the etude has involved artistic affordances based on settings in each MA, combining the three MAs in different combinations throughout a performance, improvising with corpora containing different musical material, different controller mappings, and exploring different performance gestures with the eTube and microphone. An important variable throughout this phase has been to consider what kinds of subtle interactions and body language used by human performers might be implemented as a controller mapping with the MAs.

However, rather than a direct modelling of human interaction, this was more a reflection of what we look for in an improvising partner and then searching for solutions to implement these interactions with these specific MAs, the controller, and my approach to improvising with the eTube. I also briefly discuss challenges encountered developing a new augmented instrument related to my established performance practice as a saxophonist who commissioning new works from composers. Finally, this stage established our working practice with the MAs and uncovered larger questions surrounding machine agency, musical phrases, musical memory, performance practice, and gesture (some of which I will discuss in Chapter 6), for example.

Chapter 4: Phase 2—Developing Spatialization Models

Kasey Pocius is an intermedia artist and technologist specializing in audio spatialization.[126] Pocius and I applied for a CIRMMT Student Award for the 2022–23 academic year with the research question: How can spatializing software facilitate interaction between MAs and an improviser? Specific technical details can be found in Pocius' master's thesis and a co-authored paper for the International Conference on Arts and Humanities (ICOAH), but certain details are recounted below.[127] Pocius also brought their computer performance practice to the project, a contrast with Cusson who initially engaged less with the MAs during performances leaving me to interact with the MAs via controller mappings as described in Chapter 3. Pocius actively engaged in the performance by adjusting levels, performing live processing, and interacting with the spatialization models. This improvisation scenario might be described as a trio involving the agents, Pocius, and me.

[126] "Kasey Pocius," Kasey Pocius, accessed March 23, 2024, https://www.kaseypocius.ca.

[127] Kasey Pocius, "Expanding Spatialization Tools for Various DMIs" (Master's Thesis, McGill University, 2023); Kasey Pocius, Tommy Davis, and Vincent Cusson, "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, 2024, 20–37, https://doi.org/10.17501/23572744.2023.10102.

Fig. 4.1: A workshop to record corpora material and test spatialization models with Pocius at CIRMMT.[128]

4.1 Pocius' Spatialization Models

Pocius developed two spatialization models, one for DYCI2 and Construction III, and one for Spire Muse. After Pocius designed the initial models, we met for a workshop at CIRMMT to test the spatialization models with the musical agents (MA) (see Fig. 4.1). This workshop served two purposes: 1. I improvised on the eTube to test the models in an acoustic environment, 2. Pocius recorded my improvisations, which would be used as corpora for their fixed work *eTu{d,b}e de labo #1* and *3tube* for three eTubes and electronics. The latter composition will be addressed in Section 5.3.1 below. Inspired by Bowers and Archer's infra-instrument concept, Pocius designed these models to create complex spatialization results with relatively simple interactions.[129] Rather than controlling each source individually, Pocius' models allow both them and me to interact with the MAs based on our current performance practices, a laptop and MIDI controller and the eTube and microphone respectively. The models treat the MA's total audio output as a

[128] Pocius, Davis, and Cusson, "Spatialization Models," 29.

[129] Pocius, Davis, and Cusson, 27.

larger unit that move through virtual space in cohesive ways. Please see Pocius' master's thesis for more technical details regarding their spatialization models.[130]

4.1.1 The Preset Spatialization Model: DYCI2 and Construction III

Pocius designed a preset spatilization model for DYCI2 and Construction III in Spat5, which is a software for spatializing sound in real-time.[131] This system is based on a series of twenty-four spatialization presets (see figure 4.2) that may be influenced by the improviser or the computer performer. The green spheres represent where the agent's sound sources are placed throughout a multichannel setup. It is important to note that the speakers do not move, rather the green spheres represent how the loudspeakers diffuse the sounds to create virtual movement through the performance space. Pocius uses an envelope follower on the eTube microphone which allows me to interpolate between adjacent presets. An envelope follower creates a control signal that follows the dynamics of the microphone's signal. As I play louder or softer, the envelope follower sends this fluctuating control signal to the spatialization software, where it influences the interpolation between presets.

In Fig. 4.2, the green spheres represent the sound sources that move virtually in the performance space.[132] As one follows the figure from left to right, the four individual presets show the sound sources transitioning from completely behind the performer to surrounding the performance space. Kasey controls which two adjacent presets that are interpolated between, and the sound sources fluctuate between these two presets in relation to the control signal from the envelope follower.

[130] Pocius, "Expanding Spatialization."

[131] "Spat | Ircam Forum," accessed March 3, 2024, https://forum.ircam.fr/projects/detail/spat/.

[132] Pocius, "Expanding Spatialization," 46.

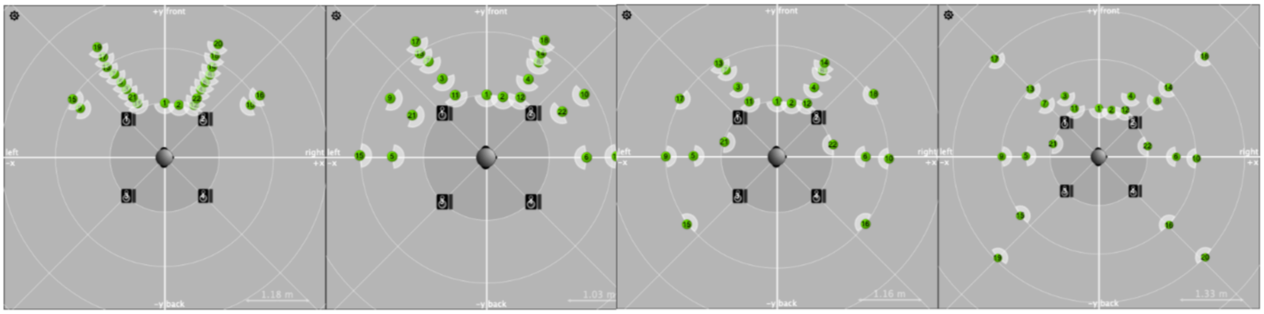

Fig. 4.2: Four spatialization presets from the DYCI2 and Construction III preset model are shown.

4.1.2 The Cluster Spatialization Model: Spire Muse

The cluster spatialization model was designed by Pocius for Spire Muse and uses Spat Revolution which is a commercially available spatialization software.[133] Their intention was to organize the outputs into one cluster which maintains a certain uniformity as it scales and rotates in reaction to the eTube's mic input or their MIDI controller (see Fig. 4.3). Their idea was that the 16 outputs could be more easily separated as they move, but the cluster would "maintain coherence as an entire entity."[134] This cluster is made of 16 audio channels from Spire Muse. The rotation of the cluster is based on the pitch chroma of the eTube's signal and the cluster's scale is based on the confidence factor of the pitch chroma analysis. Pitch chroma is a descriptor that represents tonal content in the form of pitch classes which are twelve semi-tones irrespective of octave. As I play different pitches, the pitch chroma analysis would cause the cluster to rotate around the z axis (up and down). The confidence tracker measures how confident the pitch chroma is at a given time. If I played high harmonics, the tracker was usually more confident, resulting in the cluster scale decreasing. If I played something noisier, or in the low register, the tracker was usually less confident, and the cluster size would increase. The confidence metric

[133] Pocius, "Expanding Spatialization," 47–48; "SPAT Revolution," FLUX:: Immersive, accessed March 23, 2024, https://www.flux.audio/project/spat-revolution/.

[134] Pocius, Davis, and Cusson, "Spatialization Models," 28.

would vary between pitched material and breathy or noisy contemporary techniques and is well-suited to the eTube's sonic palette and my performance style.

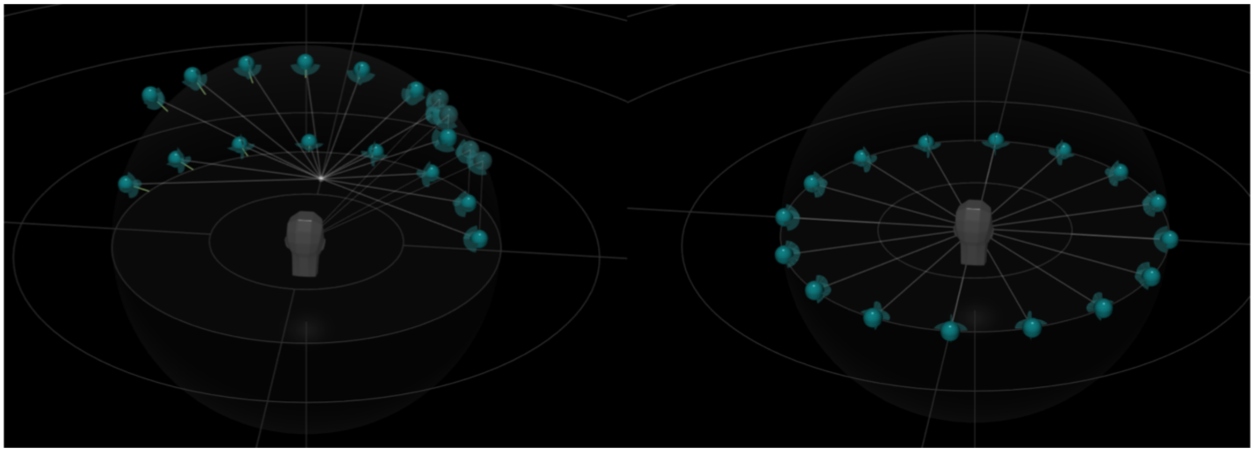

Fig. 4.3: The cluster spatialization model for Spire Muse.[135]

Slap tongue attacks are a good example of a contemporary technique that has a specific effect on the cluster model. The initial percussive attack is noisy, resulting in the pitch tracker having a low confidence. The harmonics begin to decay following the attack, and the pitch tracker is more confident as there are fewer harmonics, and it can more easily isolate specific ones. When watching the model move in real time during a slap tongue, there is a quick expansion (low confidence) following the attack, followed by the scale slowly decreasing as the pitch tracker picks up resonant harmonics (high confidence).

4.2 Case Studies

4.2.1 Cod'a 2022: *Échanges*

The first public performance using Pocius' preset spatialization model (see Section 4.1) was *Cod'a 2022: Échanges* produced by *Codes d'accès* at Eastern Bloc in Montreal on October 18,

[135] Pocius, "Expanding Spatialization," 47.

2022.[136] For this performance, Pocius worked with Cusson and me to update the *eTu{d,b}e* NIME 2022 performance patch (see Section 3.4.3) to add collaborative quadraphonic spatialization, more hands-on mixing of the agent outputs, and further control of the processing from Construction III.[137] We used three DYCI2 agents, the first of which used the water tube corpus from the NIME 2022 performance. Agent two consisted of a short low drone recorded with a Barcus Berry 4000XL contact microphone, chosen for its timbral qualities.[138] Agent three included articulated eTube harmonics, which added rhythmic variation and response.

Pocius contributed to the overall form of the improvisation in several ways. Firstly, they suggested a "dog bone" structure where all the agents and effects processing begin at full volume, quickly decay shortly after, and are slowly reintroduced slowly throughout the improvisation until all processing is gradually removed and the end features only acoustic eTube.[139] Pocius also performed with their Akai MIDImix controller near the front of the stage, in contrast with Cusson's more hands-off approach described in Chapter 3.[140] Using the MIDI controller's faders, Pocius adjusted the agent's levels and processing, shaping the overall improvisation via mixing. I interacted with the MAs and the spatialization models via the eTube's mic and controller. It was rewarding and inspiring to work with the spatialization models. It was clear from this early performance that hearing the spatialized agents clearly, and from all loudspeakers, is a challenge in performance. These issues will be discussed below in Section 4.3.

---

[136] "Cod'a 2022: Échanges," Codes d'accès, accessed February 29, 2024, https://codesdacces.org/evenement/coda-2022-echanges/; *eTu{d,b}e by Kasey Pocius and Tommy Davis – Cod'a 2022: Échanges*, YouTube video, 19:40, 2024, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

[137] Pocius, "Expanding Spatialization," 49.

[138] "Home," KMC Music, accessed March 23, 2024, https://kmcmusic.com/.

[139] Pocius, "Expanding Spatialization," 49.

[140] "MIDImix," accessed March 23, 2024, https://www.akaipro.com/midimix.

4.2.2 *Enfants, apprenez-nous à parler* by Quentin Lauvray

Enfants, apprenez-nous à parler (*Children, teach us to speak*) (2022) by Quentin Lauvray (b. 1997) is the first commissioned piece for the eTube and MAs, commissioned by Weather Vane with the support of the Canada Council for the Arts. *Enfants, apprenez-nous à parler* explores motherese and baby talk as a metaphor for the expressive but limited proto-instrument qualities of the eTube (see program note in Appendix D). The composition develops in the same way that infants are thought to learn language, first via rhythm, timbre, and melody, followed by syntax and meaning. The work's form follows this learning process and begins with simple rhythmic and melodic fragments which develop into longer phrase structures throughout. This piece features composed and improvised sections using spatialised fixed audio and five DYCI2 agents with spatialization work undertaken by Pocius. For the corpora used with the DYCI2 agents, I improvised on the eTube while listening to recordings of infant and mother communicating. Quentin manually segmented and categorized these recordings by hand based on the performance technique and sonic gesture. These recordings of categorized techniques constitute the corpora used with the five DYCI2 agents in the improvised sections.

Lauvray proposed notation solutions for the eTube since it is a limited in technique, and certain overtones are not always consistent. For this reason, Lauvray used boxed notation to specify pitch zones to be performed for most gestures, rather than writing specific pitches (see Fig. 4.4). Lauvray also considered my spatialized performance practice and gestures into the written score and indicates movements across a horizontal axis and spinning the eTube in front of the mics. Figure 4.5 exemplifies the graphic indications notated above each stave to specify certain movements with the eTube. Staff 1 indicates for me to perform a circular movement which increases in speed in front of the mic (see Fig. 4.5, staff 1). Staff 2 indicates that I should at first sweep the eTube from right to left, followed by a continued side-to-side sweeping motion which accelerates, and then begin to slow down at the end of the gesture (see Fig. 4.5, staff 2). These gestures are also related to the fixed electronics and the real-time spatialization. Lauvray proposed a microphone setup using three RE20 mics, which allowed me to perform these gestures while maintaining a consistent signal for the MAs and electronic processing. Using three RE20s also allows us to easily transition between *Enfants* and *eTu{d,b}e* without changing the setup.

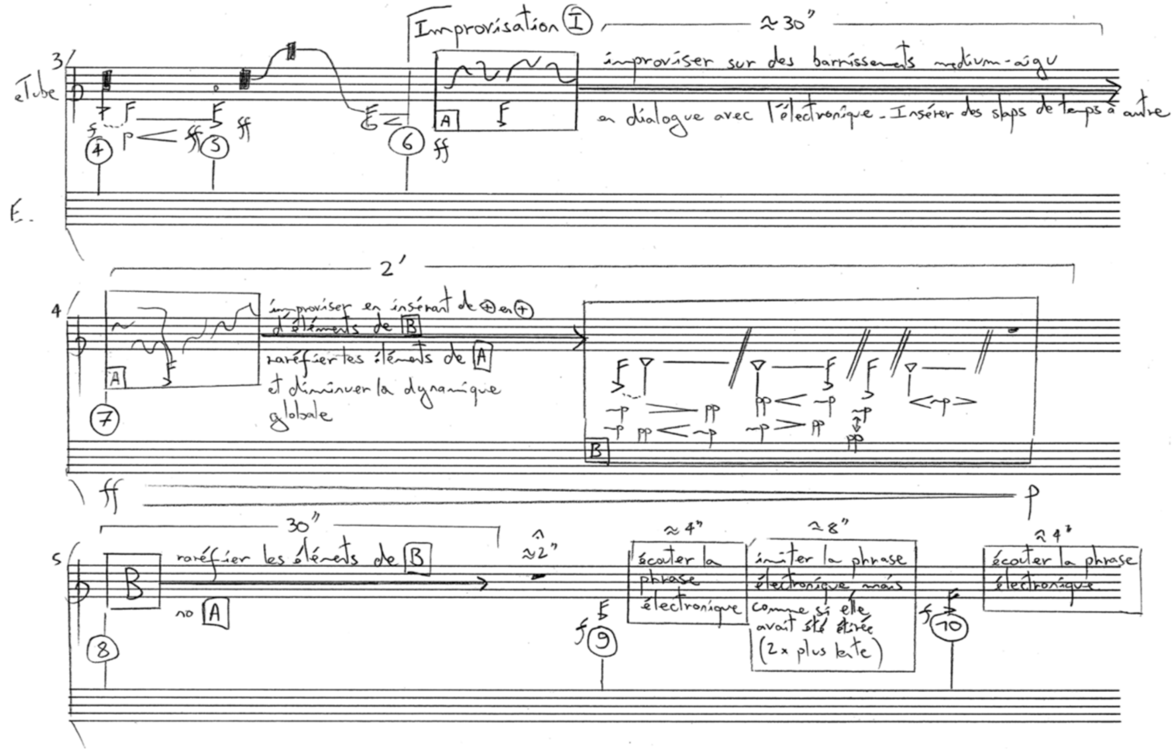

Fig. 4.4: The first improvised section of *Enfants, apprenez-nous à parler*, staves 3 to 5.

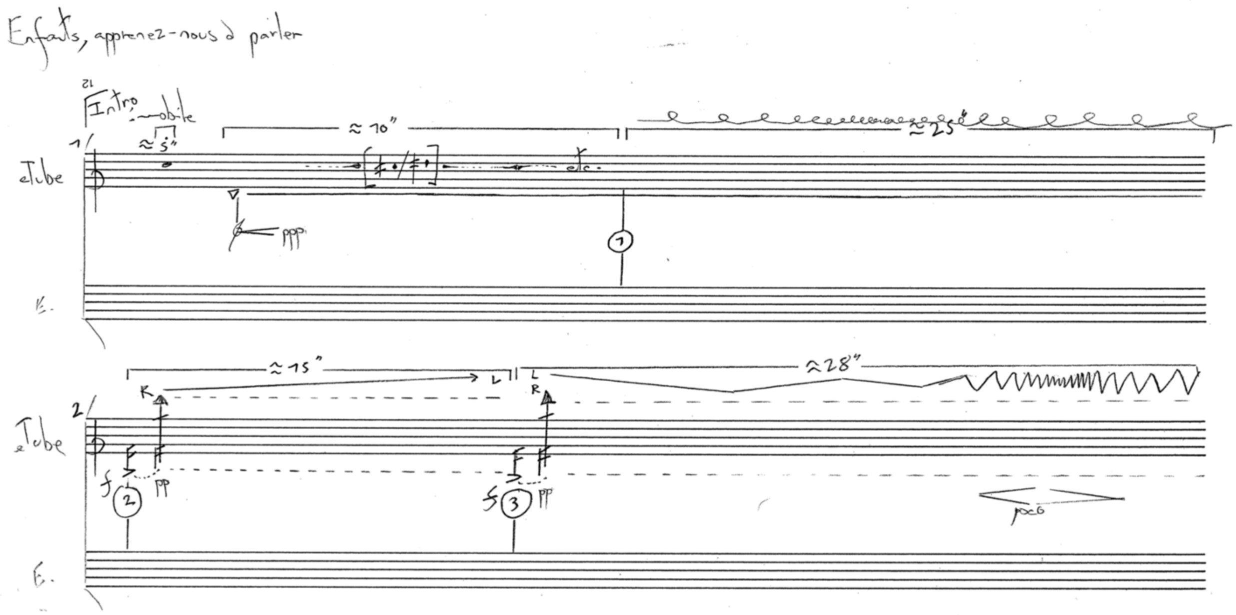

Fig. 4.5: *Enfants, apprenez-nous à parler*, staves 1 and 2.

4.2.3 Weather Vane: *Improvisation Frameworks* by Greg Bruce and Tommy Davis

As stated in the introduction, certain factors have been identified as limiting other users' adoption of new interfaces into their own performance practice. Clayton Rosa Mamedes and colleagues suggest that a lack of repertoire, dedicated instrumental technique, and DMI-specific notation has hindered wider acceptance of DMIs (digital musical instrument) in general.[141] Although these issues are varied and intricate, work by Raul Masu and colleagues suggests that the structure of the New Interfaces for Musical Expression (NIME) conference, specifically the "new" in the NIME acronym, prioritizes new technological developments rather than encouraging papers about updates and new directions in older DMI designs.[142] As Masu argues, this has led to a focus on technological developments over musical ones, and the larger risk that this "ideology of newness"[143] as Patrick Burkart claims, "mistakenly places technology, and not human agency, at the source of human history-making."[144] Considering these challenges, we made a commitment to commission works for the eTube, perform the eTube with other instruments, and work with other artists interested in performing with the eTube. Weather Vane is a Montreal-based collective comprised of saxophonist Greg Bruce, clarinetist Maryse Legault, and me.[145] Our mandate is to demonstrate a diversity of approaches to augmented instruments, highlighting the creative potential of technologically extended woodwinds.

Improvisation Frameworks (2022) is a co-improvised work by Greg Bruce and me with the support of the Canada Council for the Arts (see program note in Appendix E). This was the first collaboration with the eTube and another instrument. Bruce performed on his feedback saxophone, an augmented system he developed during his doctorate at the University of Toronto

---

141 Mamedes et al., "Composing for DMIs," 509.

142 Masu, Morreale, and Jensenius, "The O in NIME."

143 Masu, Morreale, and Jensenius.

144 Patrick Burkart, *Music and Cyberliberties*, Music Culture (Middletown: Wesleyan University Press, 2012), 121, muse.jhu.edu/book/435.

145 "Home," Greg Bruce Music, accessed February 25, 2024, https://www.gregbruce.ca; Maryse Legault, "Music," Maryse Legault, accessed February 25, 2024, https://maryselegault.com.

which generates feedback tones using personalized saxophone fingering patterns, a microphone, a guitar amp, and effects pedals.[146] This work included three DYCI2 agents trained on corpora by Bruce and me. The improvisation consisted of four sections, the first of which was for acoustic eTube and feedback saxophone only. The second section used an agent trained on Bruce performing feedback tones, and we allowed space for the agent take the lead role. The third section used my water tubes corpus. The final section used the previous two corpora, plus the Barcus Berry pedal corpus, and we built towards a final climax to end the piece.

4.2.4 *live*@CIRMMT: Improvising New Winds

Weather Vane's inaugural concert was "*live*@CIRMMT: Improvising New Winds" held at CIRMMT's Multimedia Room (MMR) at McGill University on December 9, 2022.[147] Below I will discuss the premiere performances of *Improvisation Frameworks* and *Enfants, apprenez-nous à parler*, including relevant spatialization details.

For the premiere of *Improvisation Frameworks*, Bruce and I wanted to create a trio situation including the MAs, Bruce, and me.[148] A contact mic on the eTube fed the MAs, so Greg's performance did not inform the MA's listening functions. We decided to diffuse all the MAs from a single loudspeaker and with Bruce and I set up on stage flanking either side of this loudspeaker. Pocius and the CIRMMT technicians automatized a red spotlight's intensity to the MA's output amplitude. When the MA was loud the light would be bright, and when it was soft, the light would be dim. We hoped that this would help the audience understand that we use only

---

[146] Gregory Andrew Bruce, "Feedback Saxophone: Expanding the Microphonic Process in Post-Digital Research-Creation" (doctoral dissertation, University of Toronto, 2023), https://hdl.handle.net/1807/130068.

[147] "live@CIRMMT: Improvising New Winds," CIRMMT, accessed February 28, 2024, https://www.cirmmt.org/en/events/live-cirmmt/improvising-new-winds.

[148] *Improvisation Frameworks by Greg Bruce and Tommy Davis [Premiere]*, YouTube video, 10:10, 2024, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

one agent in sections two and three, and this worked well. However, in section four of the piece, when we diffused multiple agents from the single speaker, the sounds became very dense, and it was difficult to differentiate between the various agents. Most likely, the audio would have been heard as one sound source, rather than three distinct sonic spaces. In addition, the light represented the aggregate output of all agents diffused by the loudspeaker, which again did not help audiences to distinguish between the three individual agent outputs. This also poses conceptual issues, since we curate the corpora to play specific roles, or at the very least to introduce contrasting material, in the improvisation. The lighting effect also acts to highlight the physical loudspeaker as the sounding body of the agents. Having these contrasting materials diffused from one single loudspeaker, which the spotlight signified as a single sound source, was not congruent with how we conceptualized of the agents as separate entities. Having separation between sounding bodies is important as audiences may rely on both visual and sonic cues to understand the performance environment. The physical bodies and instruments, and the space between, helps audiences to distinguish between and localize different sounds to understand the relationships between different performing bodies.

Greg and I originally conceived of a one-loudspeaker setup for prior performances in spaces with modest equipment available. Although it would have been more work to automatize three spotlights, using one loudspeaker for each agent would have resulted in a more compelling performance. This might have complicated our positioning and necessitated a wider platform stage, although rather than flanking one loudspeaker, Bruce and I could have been interspersed between the three individual loudspeakers in this setup. The significant effect of the spatial stage setup in this case exemplifies how Pocius' spatialization work would advance the eTube project both conceptually and artistically, contributing especially to the notion of the agents as separate sounding bodies (see Section 6.6 and 6.11).

For the premiere performance of *Enfants, apprenez-nous à parler*, one of the biggest challenges was hearing the electronics in the multichannel environment (see Section 4.3 for monitoring challenges).[149] The stage was placed at the far end of the hall, with one set of speakers close

---

[149] Nika et al., "DYCI2 Agents."

behind me, just above my head. This made it difficult to hear the MAs during the improvisation, depending on how the spatialization placed their output in the hall.

For this performance, the eTube controller was used to advance through presets in Lauvray's Max patch used for the fixed audio and processing. When performing mixed music with a saxophone, I often use a MIDI pedal to advance presets. A performer usually presses a MIDI pedal with their foot while playing, and this often telegraphs to the audience that a change in the electronics is about to happen. In contrast, the controller presented an integrated solution that could be activated with a subtle gesture compared to a MIDI pedal. However, there were minor delay issues with the controller's wireless signal.[150] When Cusson originally designed the controller, we had discussed the fact that there would be a short delay before the messages reached the computer. We had decided that, for the types of improvised interaction and controller mappings we were interested in, that I would seldomly need a controller input to be perfectly synchronized with the electronics. In Lauvray's piece this was the case, however. There were certain phrases where I needed to attack in unison with the electronics cued by the controller. The controller was not ideal for this situation. We were fortunate to have had significant rehearsal time in the hall prior to the concert and this was sufficient for me to be comfortable with the controller's delay. Although we used the controller in a way that it was not optimal, I was able to adapt my performance to the controller's delay to accurately align my attacks with the electronics.

4.3.3 North American Saxophone Alliance Biennial Conference

Pocius and I were accepted to perform for the North American Saxophone Alliance (NASA) Biennial Conference held from March 30–April 2, 2023, at the University of Southern Mississippi in Hattiesburg, MS. This was our first public performance with Pocius' cluster

[150] Johnty Wang, Axel Mulder, and Marcelo M. Wanderley, "Practical Considerations for MIDI over Bluetooth Low Energy as a Wireless Interface," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Porto Alegre, Brazil, 2019), 25–30.

spatialization models and the US premiere for the *eTu{d,b}e* framework. Pocius used an updated version of Spire Muse, now referred to as the Co-Creative Communication Platform (CCCP).[151] The one CCCP update I will address here is a rhythmic quantification function that was selectable on the UI. This meant that CCCP would alter the MA's audio output so that it would align with a rhythmic pulse derived from my live input. This was a saxophone conference, so we wanted to clearly showcase the instrument and the MAs. We chose to begin with a short acoustic improvisation, followed by Construction III effects, DYCI2 agents, and finally CCCP agents in the second half. Near the end of the first half, DYCI2 was outputting more soloistic material and I decided to take a supportive role, contextualizing DYCI2's output as a solo by performing rhythmic slap tongues as accompaniment. This was an exciting first opportunity to perform with Pocius' analogue synthesizer corpus which was featured in the second half with CCCP. CCCP closely matched my high overtones and textural teeth-on-reed techniques at the beginning of the second half, resulting in a complimentary dialogue between my acoustic textures and the synthesizer corpus. CCCP was outputting a similar segment which was in a clear rhythm thanks to the rhythmic quantification function. In the past, we would sometimes get tired of hearing certain segments output repeatedly during rehearsals. However, at the concert Pocius noticed that audience members were entraining with the rhythmic quantization, showing that this rhythmic and repetitive output was musically satisfying for audience members.[152] One notable distinction with Pocius' synthesizer corpus is that it had a consistent output, with little dynamic range compared to my acoustic corpora. As a result, it was challenging to balance my live sound with the CCCP agents during the performance since the output was in general, more consistent than my acoustic sound.

---

[151] Thelle and Wærstad, "Co-Creative Spaces"; "Sirnotto/Cccp: Co-Creative Communication Platform," accessed March 26, 2024, https://github.com/sirnotto/cccp.

[152] Kasey Pocius, personal communication with the author, April 2, 2023.

4.3.4 Instruments, Interfaces, Infrastructures Conference

Pocius and I were accepted to perform *eTu{d,b}e* and *Enfants, apprenez-nous à parler* at the 2023 Instruments, Interfaces, Infrastructures: Interdisciplinary Conference on Musical Media Conference (IIICON2023) hosted at Harvard University from May 11–13, 2023.[153] For this performance we used DYCI2, Construction III, and Spire Muse, including Pocius' preset and cluster spatialization models. During rehearsals, we found that reducing the segment size used by DYCI2 helped to create more varied output, which also helped to differentiate the agent from the acoustic instrument.[154] The venue supplied two AKG C414 microphones for our setup, which was a contrast to the three RE20s usually used for performances of both *Enfants* and *eTu{d,b}e*. Lauvray adapted the patch so that a third microphone source was derived from material shared by the two C414s.

In the *eTu{d,b}e* improvisation, the C414s reacted differently than the RE20s. Although Pocius and I had practiced with the C414s in the DCS prior to the conference, the MA's reactions at the performance were surprising. The levels during the concert were much higher than our sound check, so the MA's output was louder in the space overall. However, the louder speaker levels may have fed into the C414s and contributed to launching the MA's output more easily. Perhaps for this reason, I felt I did not have the same flexibility to play with the space around the microphone, as my sonic gestures would easily launch the MAs compared to our rehearsals. As a result, I played quieter overall and kept the eTube farther from the mics. I never placed the eTube directly on the microphone grill as I often would with the RE20s. The C414s did enhance the Doppler effect as I rotated the eTube back and forth in front of the mics at different rates. I played with this rhythmic aspect created by the Doppler effect and Pocius responded later by playing back recordings of this gesture using Construction III. At one point, I heard the agent output a slap tongue phrase that was panned between the back speakers. I responded with a

[153] "Conference Program," Instruments, Interfaces, Infrastructures: An Interdisciplinary Conference on Musical Media, accessed February 24, 2024, https://sites.harvard.edu/instruments-interfaces-infrastructures/.

[154] Pocius, "Expanding Spatialization," 52.

similar gesture and performed slap tongues while rotating the eTube back and forth between the two mics (see cited video at timestamp 7:30–9:10).[155] This performance experience solidified our understanding of the microphone as an integral part in the eTube system and performance practice. It also reinforced the need for us to have a consistent mic setup for rehearsals and performances.

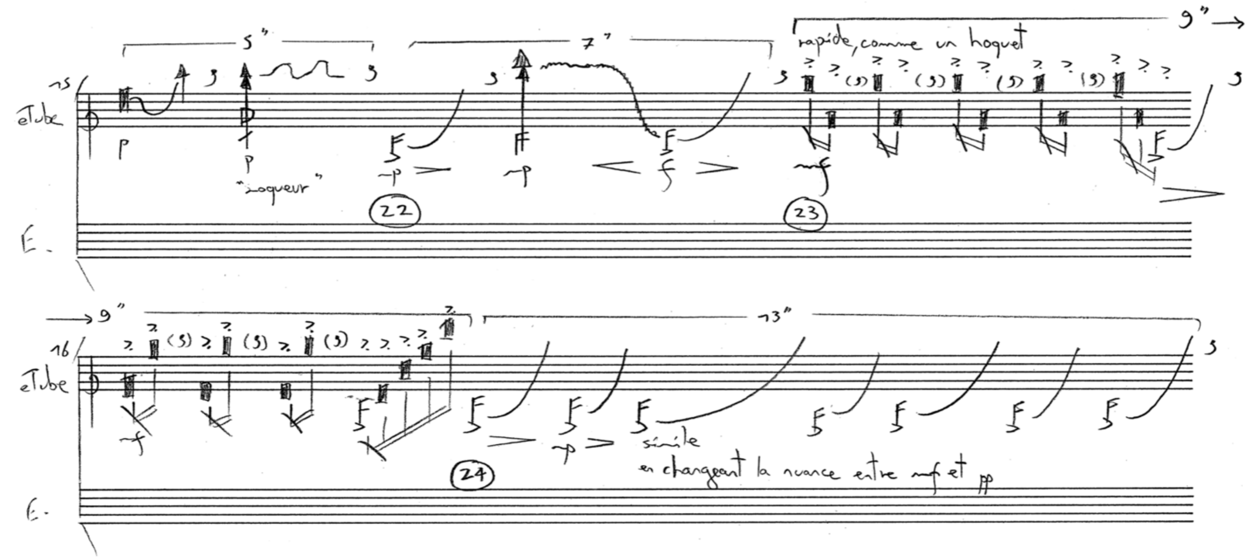

Fig. 4.6: *Enfants, apprenez-nous à parler*, staves 15 and 16.

I also performed Lauvray's *Enfants, apprenez-nous à parler* on this program, and Lauvray had updated the electronics for staves 15 and 16 since the *live*@CIRMMT performance. Lauvray composed a delay effect which cascades throughout the speakers following my initial input (see cue 23 in Fig. 4.6). Cue number 23 launched the echo effect in the electronics. I needed to align the written rhythmic gestures with the fourth iteration of the spatialized delay in the electronics. In performance, I would often not clearly hear the delay effect as it passed through the rear speakers, which made it more difficult to play this section rhythmically accurate and line up the following attacks with the cascading delay. During a recording session at a later date, I could hear this effect in my in-ear monitors, which helped me to coordinate the gesture in time with the electronics.

[155] *eTu{d,b}e by Kasey Pocius and Tommy Davis – IIICON 2023*, YouTube video, 9:56, 2023, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

4.3 Can I Hear You? Monitoring in Multichannel

Performing in an immersive multichannel environment can be challenging for musicians for many reasons. Firstly, the ideal listening position for a multichannel setup is often in the centre of the hall, usually far from the onstage performer. Secondly, the performer will hear the closer loudspeakers more easily, and this will mask sounds diffused elsewhere in the space. I will often hear one set of speakers more clearly, since they are physically closer, and more distant speakers will be difficult to hear. Thirdly, one must consider that bone conduction transfers sound from the woodwind mouthpiece and vocalizations from the performer to the inner ear via the teeth and skull.[156] As a result, the electronics may be masked by bone conduction or bodily resonance when a performer is playing at louder dynamics, contributing to monitoring challenges. This can also be an impetus to stop playing and leave space to listen to the MAs, as one would in an improvisation with other humans, although this does not help if one cannot hear details diffused from the hall's opposite end.

In-ear monitors may help a performer hear the electronics since the monitors block out ambient noises and the mix may be adjusted for their needs.[157] However, bone conduction for wind instruments is increased when the ear is occluded, so sound from bone conduction may mask softer electronic sounds in the monitor mix, especially if the monitors are set to a lower volume for aural health.[158] We have successfully used binaural headphone monitoring during studio recording where we have the necessary time for setup, and this provides a clearer image of the

---

[156] Jonas Braasch, *Hyper-Specializing in Saxophone Using Acoustical Insight and Deep Listening Skills*, vol. Vol. 6, Current Research in Systematic Musicology (Cham, Switzerland: Springer, 2019), 35, 43, https://doi.org/10.1007/978-3-030-15046-4.

[157] Jeremy Federman and Todd Ricketts, "Preferred and Minimum Acceptable Listening Levels for Musicians While Using Floor and In-Ear Monitors," *Journal of Speech, Language, and Hearing Research* 51, no. 1 (2008): 147–59.

[158] Robert Albrecht et al., "Electronic Hearing Protection for Musicians," in *Proceedings of the 14th Sound and Music Computing Conference*, 2017, 306–13, https://urn.fi/URN:NBN:fi:aalto-201708036338.

MAs spatialized gestures than a stereo mix.[159] However, the hall's acoustic reflections and interactions are not respected in binaural mixes which limits my ability to engage with the hall's acoustics, diminishing the real-time connection between me and the space.[160] Since directionality, movement, and interaction with acoustic space are important to my eTube performance practice, in light of some benefits mentioned above, we limit in-ear monitoring in live performances. In addition, our sound checks are often limited in duration and a binaural mix currently requires too much technical overhead to warrant regular performance use.

One of Pocius' solutions to improve monitoring for me is to place the MAs in the speakers at the back of the hall.[161] If the speakers are placed correctly, I will hear the MAs via the back speakers since they project directly towards the stage, whereas the front speakers are often in front of me and face away from the stage. This decision affects the mix and the audience's experience since the agents are consistently diffused out of the back speakers. However, I have found that this is a relatively simple and effective solution, with the benefits outweighing the artistic drawbacks. This approach depends on the hall's acoustics and speaker placement, however.

4.4 Chapter 4 Conclusion

Chapter 4 describes the spatialization models designed by Kasey Pocius which allows me to influence the MA's spatiliazation in real-time via the existing microphone setup already in use. Multichannel performance environments pose monitoring challenges as a performer, and although I could use in-ear monitors to receive a controlled mix, this would not represent the interaction of my live sound in the hall, which is an important aspect of my performance

[159] Valentin Bauer et al., "Binaural Headphone Monitoring to Enhance Musicians' Immersion in Performance," in *Advances in Fundamental and Applied Research on Spatial Audio*, ed. Brian F. G. Katz and Piotr Majdak (Rijeka: IntechOpen, 2022), 193–219, https://doi.org/10.5772/intechopen.104845.

[160] Bauer et al.

[161] Pocius, Davis, and Cusson, "Spatialization Models," 34.

practice. Bone conduction headphones transmit sound via the skull, leaving the ears open to the acoustic space. These new technologies may be a viable option in the future that would allow a cleaner overall mix, while still allowing me to hear my acoustic sound interact with the space.

It is important to the eTube team that other performers, programmers, or composers become involved with our project to share our work and extend the longevity of the eTube project (see Section 1.1). To encourage other performers who might be interested in performing the eTube, I commissioned Quentin Lauvray to write *Enfants, apprenez-nous à parler*, the first composed work for the eTube, electronics, and DYCI2 agents. Building on the work established in Phase 1, this second phase focused on performing and presenting the eTube at conferences, festivals, and concerts, which is evidenced by the Case Studies presented. Both Cusson and Pocius have focused on maintaining adaptable and flexible software and hardware that functions in different performance environments, and this was exemplified in the diverse spaces that we have successfully performed in. This is also a result of working with highly competent and dedicated collaborators. We have also received invaluable feedback and questions from other artists and researchers from the various communities we performed in.

Chapter 5: Phase 3—Updating the eTube Controller Hardware and Firmware

Maxwell Gentili-Morin is an electrical and software developer.[162] We were awarded a CIRMMT Student Award for the 2023–24 academic year to update the eTube's controller hardware and firmware, and to develop haptic feedback for additional communication with the musical agents (MA). Since 2022, Cusson has continued updating and repairing the controller, and Cusson and Pocius have made their own software updates. Phase 3 also included a collaboration with the Weather Vane collective, resulting in new eTube models, mouthpiece adapters, and performance practice.

5.1 Resin Mouthpiece Adapters and Key Caps

I will outline two hardware updates Cusson made between 2022–24. Firstly, Cusson printed mouthpiece adapters with a 3D printer using resin (see Fig. 5.1). Rather than using electrical tape to fit the mouthpiece, Cusson had an instrument repair shop add cork to the mouthpiece fitting, like a saxophone neck. Secondly, Cusson designed and 3D printed resin keycaps in different sizes to suit different preferences (see Fig. 5.2). The 3D printed resin parts are more durable and professional looking than the original plastic parts used for the beta version. As a result, they help to make the instrument more reliable during travel, rehearsal, and performances. However, resin 3D-printers are not as common as plastic 3D-printers, and these resin components are not as accessible as the off-the-shelf parts we used for the beta version during the first two years.

[162] "Maxwell Gentili-Morin – IDMIL," accessed March 26, 2024, https://www.idmil.org/people/maxwell-gentili-morin/.

Fig. 5.1: Cusson's 3D-printed resin mouthpiece adapter with cork.

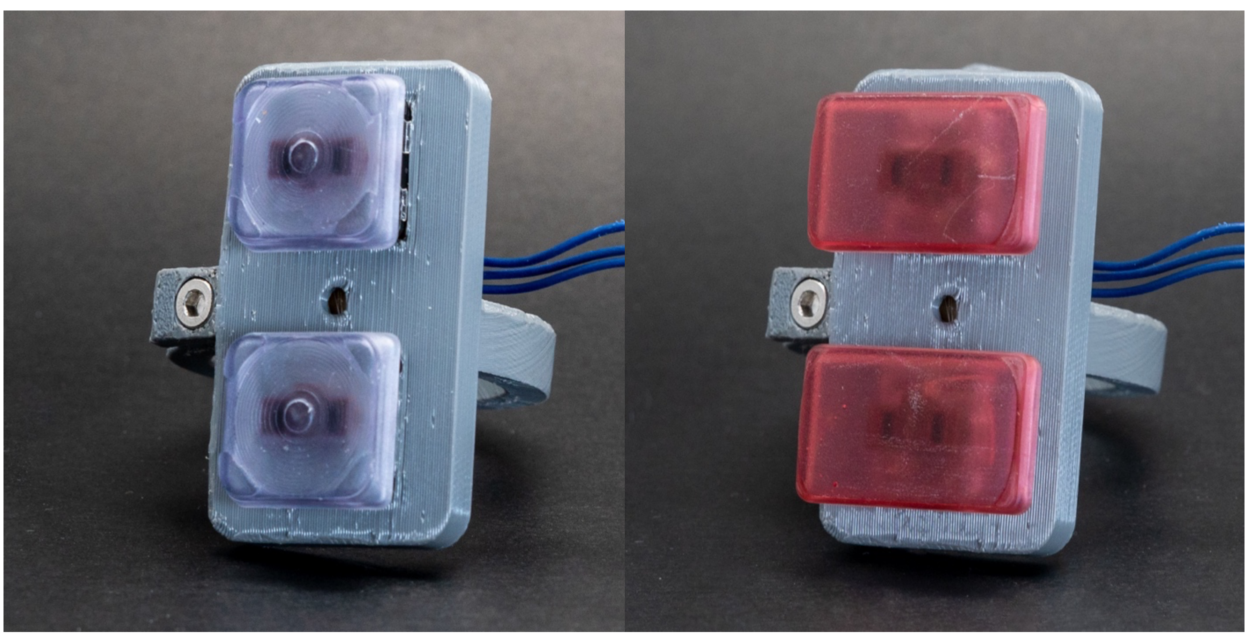

Fig. 5.2: Cusson's 3D-printed resin key caps in two different sizes.

For the *3tube* project with Pocius and the Weather Vane collective (see Section 5.3), Gentili-Morin designed and 3D printed bass clarinet and B♭ clarinet mouthpiece adapters, and Cusson designed and printed a tenor saxophone adapter. Clarinet mouthpieces have a cork which must fit into adapter, unlike the saxophone which fits over the mouthpiece fitting (see Fig. 5.3). Rather than using electrical tape on the clarinet mouthpiece cork, we used polytetrafluoroethylene (PTFE) tape to protect the cork and to maintain a seal with the adapter. The bass clarinet eTube was performed with a 2.54 cm diameter tube, like the baritone and tenor saxophone versions. The B♭ clarinet mouthpiece was used on a 1.27 cm diameter tube.

For the tenor saxophone adapter, the mouthpiece fitting needed to be smaller to accommodate the smaller tenor saxophone mouthpiece, while leaving the tube fitting the same size to fit a 2.54 cm diameter tube. In Cusson's parametric design, a change to the mouthpiece fitting on top would proportionally adjust the tube fitting on the bottom. This is one circumstance where the parametric design was not ideal, and Cusson needed to manually preserve the dimensions of the tube fitting, while making the mouthpiece fitting smaller.

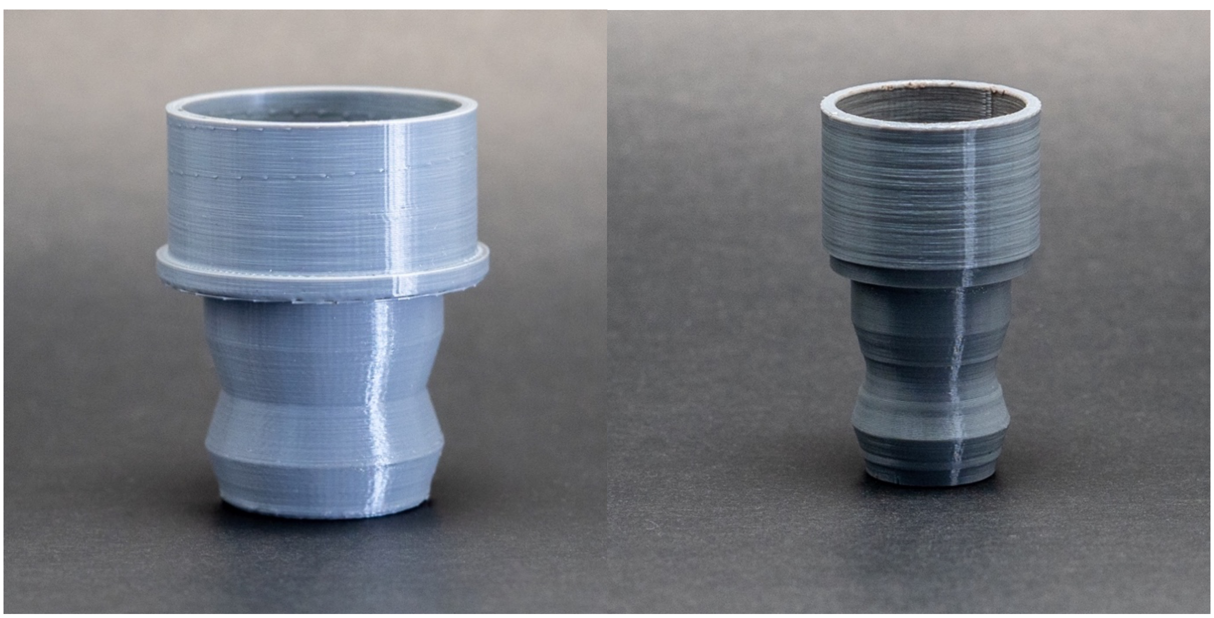

Fig. 5.3: The 3D-printed bass clarinet mouthpiece adapter (left) and the B♭ clarinet mouthpiece adapter (right).

5.2 eTube Controller Updates

As part of a CIRMMT Student Award (2023–24) Maxwell Gentili-Morin and I have worked in collaboration with Vincent Cusson to update the original eTube controller. Over the past two years as I undertook typical practice and rehearsal schedules leading up to concerts, certain parts were less reliable, and necessary improvements were made obvious. Travelling on exchange to the Metacreation lab and performing concerts around North America clearly demonstrated a need for a more robust design, and one with components that could be unplugged and disconnected from the tube during transportation. For example, the small wires connecting the buttons with the ESP32 board would often snap where the wires left the jacket, and they would need to be stripped and soldered again. The wires would sometimes break during rehearsals, but more often during travel, when I would remove the controller and buttons from the tube. The red wire broke off at the jacket while travelling to Vancouver for a research exchange at Simon Fraser University's Metacreation Lab (see Fig. 5.4, left). I broke multiple wires while sliding the battery compartment and the buttons off the eTube (see Fig. 5.4, right).

For the new controller design, Gentili-Morin designed and fabricated printed circuit boards (PCB) at The Gearbox, a makerspace at McGill's Department of Physics (see Fig. 5.5, left).[163] The PCB is below the two keyboard switches, and allows the switches to be hot-swappable, which means that these switches may be snapped in and out of place with no tools or wires to disconnect. Maxwell also made a perforated circuit board. The Tiny Pico ESP32 board is soldered to the perforated board, and it also contains the wiring for the battery and the Japan Solderless Terminal-type (JST) connector that connects the PCB and switches via the blue wires (see Fig. 5.5, right).[164] The JST connectors and wires are more durable than the original design and may be unplugged easily, allowing the controller to be removed from the eTube safely.

[163] "Physics Makerspace," accessed March 22, 2024, https://makerspace.physics.mcgill.ca/printers.html.

[164] JST is also the common name used for this type of connector, whether or not manufactured by Japanese Solderless Terminal, see Matthew Millman, "Matt's Tech Pages," *Common JST Connector Types* (blog), January 7, 2021, https://www.mattmillman.com/info/crimpconnectors/common-jst-connector-types/.

Maxwell also updated the Tiny Pico firmware to be compatible with various ESP32 microcontrollers. See Fig. 5.6 for the gamma eTube controller, and Fig. 5.7 for the eTube with updated resin mouthpiece adapter and gamma controller.

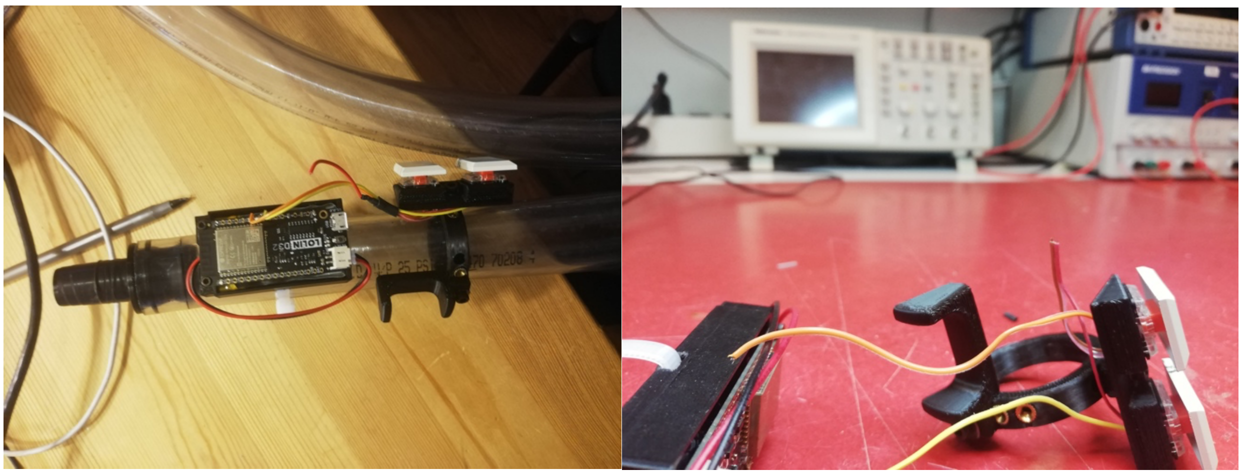

Fig. 5.4: I often broken the eTube wires during transportation.

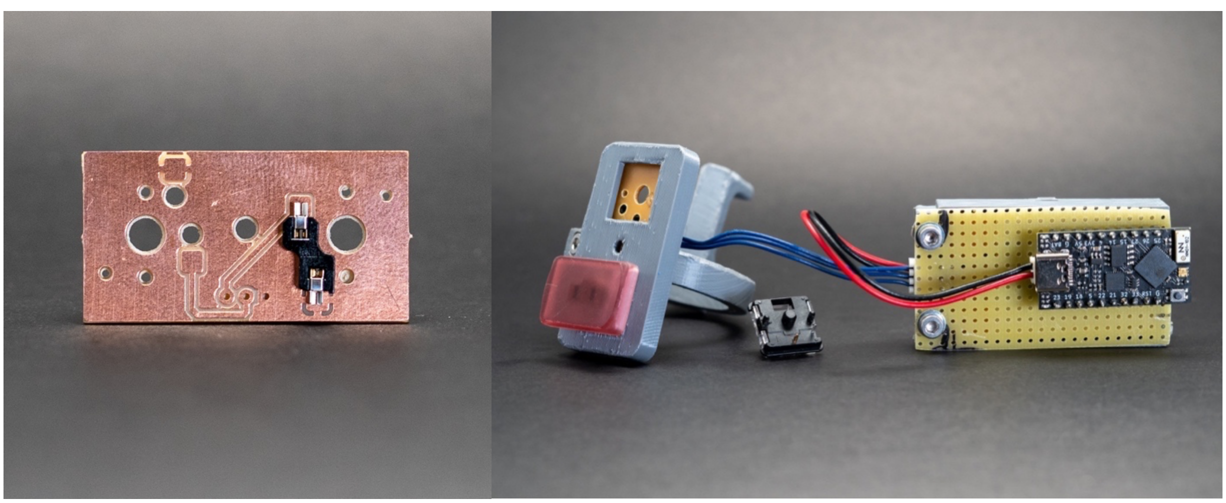

Fig. 5.5: The custom-printed PCB (left) and Gentili-Morin’s updated controller (right).

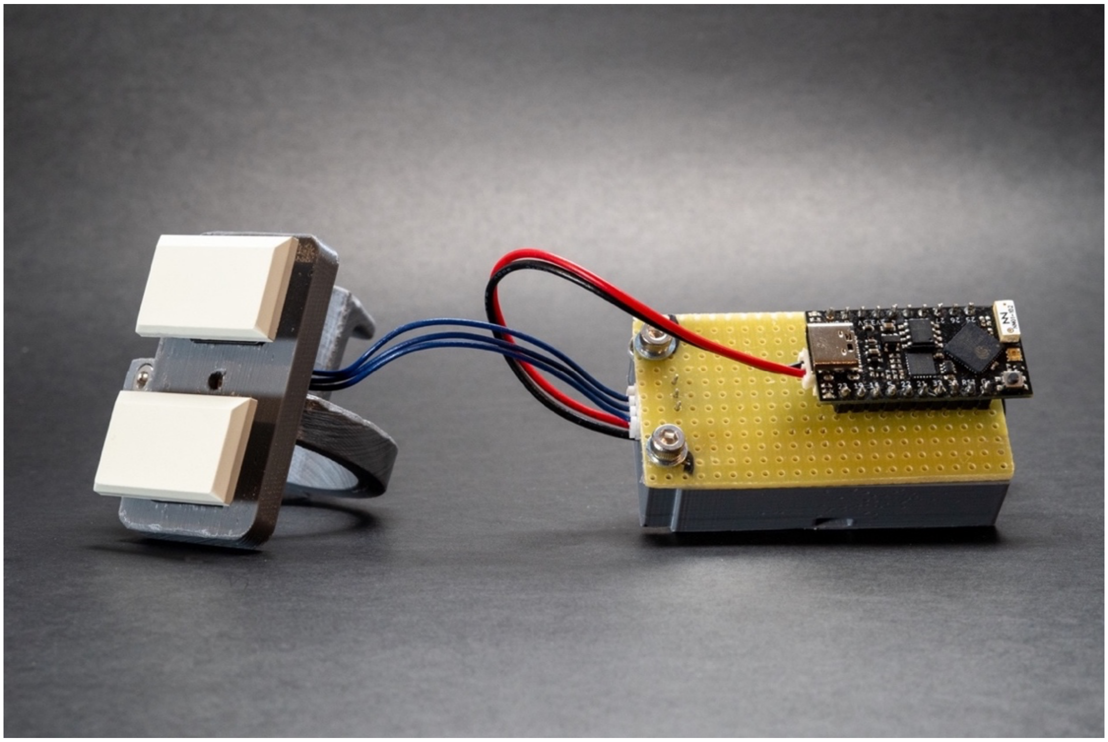

Fig. 5.6: The gamma eTube controller designed by Gentili-Morin and Cusson.

Fig. 5.7: The eTube gamma version with baritone saxophone mouthpiece.

We are planning to implement haptic feedback on the eTube controller to add an additional communication layer between the MAs and performer, providing me with more insight into the MA's ongoing processes. We decided to use haptic feedback rather than a screen since small haptic sensors may be integrated directly on the instrument. In the past, I found that having a large screen onstage was often distracting and necessitated more setup. In performance, I found myself looking at the screen when it was not necessary. I also felt that the audience would notice my eyes glued to the screen, and that this would detract from the interactive environment we were trying to create with the agents. Our plan is to use the tactile sensors to send different types of pulses such as one short, two short, or long pulses with different intensities. This feedback may be used to communicate details such as the confidence of the MA's analysis of my input, or to synchronize attacks, for example. At the time of writing, we are considering various sensors, materials, and software, and have no prototype yet. Our next steps will be to test salience, or how differentiated the pulses need to be for me to distinguish between them when performing. We will also consider how to integrate these new tactile sensors on the eTube, including battery power and cabling.

5.3 Case Studies

5.3.1 Weather Vane Collaboration: *3tube* by Kasey Pocius

3tube (2023) for three eTubes, electronics, and musical agents by Kasey Pocius (b. 1998) that was co-improvised with Greg Bruce, Maryse Legault, and me and supported by the Canada Council for the Arts. This 15-minute work is for three eTubes, spatialised electronics, and musical agents (see program note in Appendix F). The electronics part is composed of field recordings taken by Pocius and additional processing, both of which are spatialised using software. This is the first composition using the new tenor saxophone, bass clarinet, and B♭ clarinet mouthpiece adapters. Gentili-Morin assembled the new controllers for Bruce and Legault, including 3D printing parts and creating the PCBs and perforated boards. Greg and Legault were learning the eTubes at the same time the work was being co-improvised, and Gentili-Morin and Cusson were updating eTube components based on Bruce and Legault's feedback and needs. Below I will describe details related to these three avenues which occurred overlapped during many weeks and collaborative workshops.

DYCI2 had recently been updated by a new version called dicy2.[165] There were significant updates to the MAs, but for the purposes of this document, dicy2's overall operation is similar to DYCI2, and I will focus on how Pocius used dicy2 and their co-improvised creation process. Pocius used dicy2, CCCP, and Construction III along with their preset spatialization models with new presets from those described in Chapter 4. In the early stages, Pocius held workshops with each performer individually, where Greg, Legault, and I would improvise with the MAs and my existing corpora. Pocius also recorded each performer improvising on the eTube, and these recordings would be used for corpora in the piece. The creation process involved co-improvising the piece, where each performer was involved in contributing to the composition process. This co-improvised process was somewhat stunted because Gentili-Morin and Cusson were developing the new controller at the same time. In the early stages, there was only one working eTube controller, which limited our ability to test certain details together. Pocius also catered the corpora material to each performer's interests and practice. They chose more timbral audio for Legault's corpora, more rhythmic and melodic material for Bruce, and more contemporary technique-based audio for my corpora.

Pocius adapted dicy2 so that following an output from one of the three agents, the listening module would randomly choose one of the three mic inputs. The MA would listen to the assigned mic until the MA's output was launched, and then it would be randomly assigned to one of the three mics again. This was one way to allow each of us to interact with the different corpora loaded into the three dicy2 agents. If all MAs listened to all three mics, there would have been much more output from each agent as the combined input from all three mics would be much higher. This was Pocius' solution to organize the MA's output, so that the performers would not constantly launch agent outputs. Pocius had an override button which would force the MA to randomly choose another mic. This was primarily for the section where Legault would

[165] Nika et al., "Dicy2 for Max."

move far from her mic with no ability to launch an agent, to avoid multiple MAs all waiting for an input from her mic and resulting in a significant break in the MA's output.

In the piece's second section, Pocius instructed each improviser to record a motive performed by the other two improvisers. Cusson created a Max patch with multiple recording buffers, allowing each performer to record the other two performers via their respective mics. The controllers were programmed so a double click of one button would start a recording, and a double click of the same button would stop it. For example, if I clicked button 1 twice, this would start a recording from Legault's mic, and a double click of button 2 would start a recording from Bruce's mic. These recordings were saved in the patch until near the last section of the piece, when Pocius would manually diffuse the recordings, in addition to the ongoing fixed audio and MAs. Pocius' interest was to expand my practice using the controller to record and diffuse live material with Construction III but expanding the practice to a communal technique. In addition, they were interested in researching the interactive and formal possibilities when using multiple controllers. In a similar way as the agents utilize our recorded corpora, Pocius improvised in real-time by outputting our live recordings from earlier in the piece, interacting with notions of motive, memory, and reproduction.

Pocius' *3tube* score includes three elements including a spectrogram, a waveform, and notation for the three eTubes (see Fig. 5.8). The top layer is a spectrogram of the fixed electronics part created in the open-source program Sonic Visualiser.[166] Both the spectrogram and the waveform representations specific elements of the fixed electronics part. The spectral analysis in the spectrogram shows how the energy levels of frequencies vary over time but lacks temporal precision. In contrast, the waveform shows precise onsets in time, but lacks detailed information about the spectrum. Pocius has included both representations to help performers form a clearer understanding of the fixed electronics part. Pocius also uses intuitive notation for the gestures which were inspired by Lauvray's *Enfants, apprenez-nous à parler* score. These gestures were

[166] Chris Cannam, Christian Landone, and Mark Sandler, "Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files," in *Proceedings of the ACM Multimedia 2010 International Conference* (Firenze, Italy, 2010), 1467–68.

the foundation for a co-improvised process whereby Bruce, Legault and I workshopped with Pocius to developed sonic and physical gestures that each performer would explore for a specific section. As we worked together, our understanding of the piece, and our improvised performance adapted as this process evolved during the workshops.

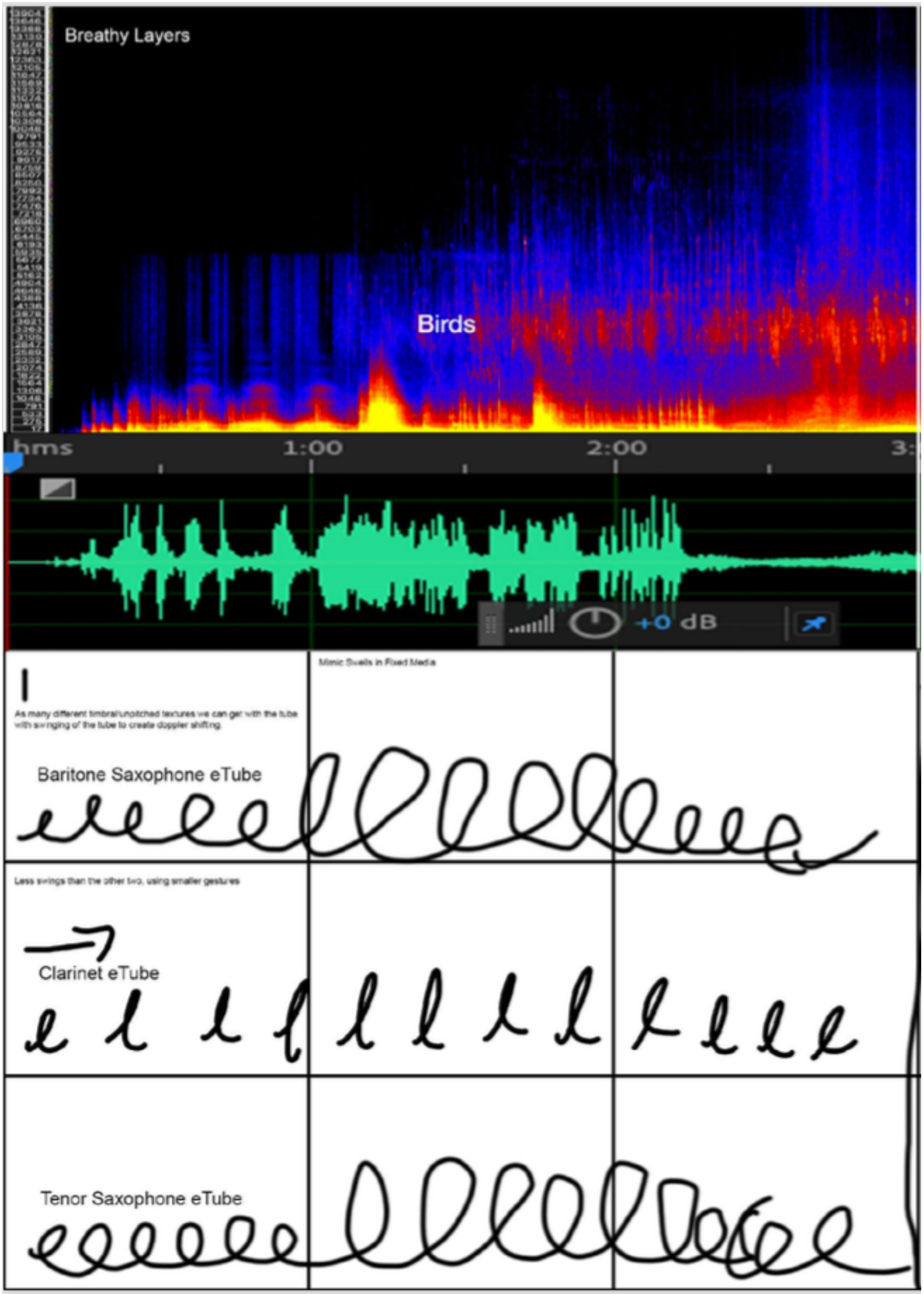

Fig. 5.8 An excerpt of section 1 from Pocius' *3tube* score.

Bruce contributed a new controller mapping during the *3tube* collaboration (see Fig. 5.9). Bruce and Legault had the same mappings for their respective controllers, with the ability to record the other two artists, but not themselves. In a workshop with Pocius, Bruce wanted to override the effects processing from Construction III so he could better blend his amplified acoustic sound with the fixed electronics part. Pocius programmed a controller action to allow the performers to override Construction III's effects so only the eTube sound and reverberation would be diffused, with no additional effects.

Controller Action	**Intended Interaction**
Single click button 1	On/off Construction III effects (reverberation always on)
Double click button 1	Start/stop recording Bruce
Double click button 2	Start/stop recording Legault

Fig. 5.9: The author's controller mappings for *3tube*.

Bruce and Legault contributed new eTube performance gestures while co-improvising *3tube*. These new additions were partially due to their individual practices, but also due to physical differences between the performers, and material differences between the tubes. I had been playing with the same eTube for years, and upon buying new PVC for Bruce and Legault's models, we found the material was much stiffer than my current eTube. As discussed in Section 3.3.4 the PVC has a natural curve from storage and transport, and the added stiffness made it less comfortable for Bruce and Legault to use the thumb rest to support the eTube. Bruce's hands are larger than mine, and when using the thumb rest to support the eTube, his fingers extended beyond the buttons. It was not comfortable for him to press the buttons in this position. Rather than using the thumb rest, Bruce braced his thumb against the bottom of the mouthpiece for support, distancing his index and middle fingers from the eTube, and allowing him to engage the buttons more comfortably.

The new stiffer eTubes were also 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 better suited for the 15-minute *3tube* performance. Notable gestural additions include Bruce abruptly pressing the sounding end of the eTube against his body to create a sharp cut-off. Although not related directly to eTube performance, Bruce chose not to perform on the eTube and sang directly into the mic. Legault interacted more with space in and around her body by interacting with the tube, but also the performance space by moving around onstage. She would wrap the eTube around her body in various orientations while playing, slowly morphing her relationship to the eTube and mic throughout the piece. Legault would move away from her mic, introducing space between her and the agent's "ear." She would also kneel on the ground in front of the mic, performing from an intimate proximity. To facilitate Legault orienting the eTube on her body in different positions more fluidly, Cusson designed and 3D-printed rings to be placed near opposite ends of the eTube (see Fig. 5.10). Legault was able to secure the instrument with one hand by hooking a finger through two different rings.

Fig. 5.10: Cusson's 3D-printed eTube ring from resin.

5.3.2 Codes d'accès: Melting Links

3tube was premiered for "melting links," co-produced by *Codes d'accès*, *le Vivier*, and CIRMMT in the MMR at McGill University on November 30, 2023.[167] This was the first performance with multiple eTube performers, the premiere of the new tenor saxophone and bass clarinet eTubes, and the first use of the B♭ clarinet mouthpiece adapter.[168] You can see the different approaches to eTube performance from the three artists. Bruce uses his voice off the eTube and sings into the mic. From timestamp 7:35 to 10:05 of the cited video, Bruce takes the role of soloist using isolated vibrational modes and rhythmic motives, with Legault and I following with accompaniment figures in the latter part. Between 8:30 to 10:35, Legault ventures away from her mic, moving through the space and away from the agent's "ear." She also interacts with the eTube around her body, while kneeling and investigating the mic from a close proximity (see 12:05–15:00). Throughout the second section (see timestamp 3:05–5:55), Bruce, Legault, and I are recording each other using the controller functions. Near the end of the piece, Kasey diffuses these motifs in alternation, bringing back the live recordings we made of each other in the first part of the piece, and signaling the final section (see 12:00–13:00).

5.4 Chapter 5 Conclusion

Chapter 5 outlined the hardware, software, and firmware updates made to the project during the 2023–24 academic year with a focus on the collaboration between Gentili-Morin, Cusson, and me to update the eTube's controller. The problematic wires and soldered joints have been replaced with JST connectors and more robust wires, a smaller ESP32 Tiny Pico board and battery, and a custom-made perforated board and PCB have been designed and manufactured by Gentili-Morin. Cusson continues to make regular improvements such as a mouthpiece adapter

[167] "melting links," Codes d'accès, accessed February 29, 2024, https://codesdacces.org/evenement/melting-links/.

[168] https://youtu.be/O37SMEwaIRo

and key caps printed in resin which have helped to make the eTube more professional and rugged. These improvements were implemented following two years of travel and performance with the eTube, which clearly demonstrated necessary updates to improve robustness, and has resulted in the gamma version of the eTube.

A collaboration with the Weather Vane Collective spurred the development of additional eTube controllers and new mouthpiece adapters for tenor saxophone, bass clarinet, and B♭ clarinet. The collaboration culminated in the creation of the second commission for the eTube, *3tube* by Kasey Pocius co-improvised by Greg Bruce, Tommy Davis and Maryse Legault. Throughout this process, Pocius developed new interactive strategies for routing mic signals to the MA's and integrated new controller mappings to allow performers to record each other. Bruce and Legault contributed new eTube performance gestures and practice, and Pocius developed notation to reflect these developments. Since we were designing and building the new eTube controller and mouthpiece adapters in tandem with the *3tube* workshops, this complicated the collaborative creation process since Pocius was only able to work with one performer at a time until we had built the additional controllers. This collaboration also fulfilled the commitment to involve other performers working with the eTube, and to commission works for the instrument, as a way to share our work more broadly and extend the longevity of the instrument.

The past three chapters have outlined the state of the project, developments, and performance case studies over the past three years. The following chapter will focus on a philosophical discussion related to human and computer agency, a discussion of machine improvisation, how we have come to consider the MAs as co-collaborative partners, and ethical considerations. I will also refer to specific details in earlier sections and performance media throughout.

Chapter 6: Philosophical Considerations—Do Musical Agents Improvise?

6.1 Introduction

In the preceding chapters, I have presented the eTube project, including technological and conceptual approaches to working with musical agents. Inherent in this discussion are certain assumptions that allowed us to engage with these technologies. In the early stages, Cusson, Pocius, and I were focused on learning about and developing the musical agents (MA) for performances. At the same time, I consulted literature on relevant subjects, performed and presented the project, and received feedback from eTube team members, collaborators, academics, and audiences. As the project advanced, I felt the need to revisit some of these initial assumptions and to investigate my relationship with the MAs, the eTube, other collaborators, and audiences.

This reflective and research-based process has been both challenging and rewarding. This dissertation is by no means a conclusive discussion on musical agents or our project. However, solidifying a philosophical perspective has helped to clarify the project artistically and conceptually, and I anticipate that it will continue to focus future artistic directions. In addition, I have grown as an artist and researcher through this process whereby I have thought more deeply about the philosophical and ethical concerns related to the project. I am indebted to the eTube team and other collaborators for their insights and support throughout this process, especially Cusson and Pocius' participation in a reading group led by Professor Eric Lewis during the 2023–24 academic year, which has informed many of the discussion and examples presented throughout.

The following section outlines philosophical positions which we build upon to make sense of how we interact with musical agents. We have chosen to describe our interaction with MAs in the ways described below because this helps us to make sense of the music that we make with them. I am interested in describing a heuristic for considering agency in the MAs through improvised performance. These descriptions will be specific to the MAs used with *eTu{d,b}e*, outlining affordances and constraints, while also referencing other programmers and artists who have contributed important developments related to these topics. The account below begins with a discussion of relevant details regarding MA improvisation and interaction. I will then follow Eric Lewis' chapter on musical agents which builds upon Daniel Dennett's intentional stance, Jerrold Levinson's musical personae theory, Theory of Mind (ToM), and fictional characters research.[169] In contrast to Eric Lewis, who focuses on George E. Lewis' MA Voyager and larger questions regarding whether agents may improvise in an Afrological way, I will describe how these concepts play into the development and implementation of our artistic practice and the *eTu{d,b}e* framework.[170] I will then discuss future work related to MA feedback in performance before addressing ethical concerns surrounding the project and interactive technologies more generally.

6.2 Improvising with Musical Agents

Referring to a performer's expectations when improvising with a MA, Oliver Bown and colleagues state that "the claim of human-like abilities is here an obstacle to achieving the best interaction."[171] Expecting that the agents will have abilities equal to a human performer often leads to expectations that cannot be met by the agents, such as nuanced communication, musical development, or dynamics, all musical characteristics that are expected from a human performer. At the same time, if an improviser constantly dominates over a system, they will not leave space for potential interaction. It is on the continuum between these two extremes that I have found the most rewarding and musical interaction with the MAs.

---

[169] Lewis, *Intents and Purposes*, 57–102.

[170] Lewis, "Too Many Notes"; Lewis, "Improvised Music After 1950."

[171] Bown et al., "The Musical Metacreation Weekend," 31.

As suggested in Section 1.2, David Borgo posits that developing a personal or group style in improvisation is part of establishing one's own expertise.[172] Within the context of established groups that practice improvisation, artists might intimately know their fellow improviser's sonic ideas and musical preferences. Improvising over many years with Nick Zoulek has led us to develop what might be considered as improvisations with specific sections and directions, and ones that we might even give names to and perform specific instances of for concerts. Over multiple performances we have explored certain musical ideas, which then suggest specific sections or ideas that coalesces over multiple improvised performances. Although the form and actual material will always be different, there is a shared road map, and an understanding of where the other improviser is headed in the context of the shared performance that constitutes a form that has been co-created together.

In terms of the MA's sonic material, my relationship to the MA is similar to the above, but distinctly different. I know the material that is in the corpora because I recorded it (except for the synthesizer corpus used by Pocius). I know this material even more intimately than I would the improvisations of a close collaborator, I have an embodied and historical relation to the sound files. However, I do not know exactly which segments, and in what order, the agent will recombine and output the corpus material. From experience with DYCI2's smaller corpora, I may have some idea of which segments might be output more than others, and certain phrase combinations that might be prioritized by the agents when I perform specific techniques. This has been more difficult to judge with Spire Muse as the corpus is larger, and often contains more diverse sounds.

Another distinction from the human-human situation above is that the MA cannot communicate gestures or body language, which is an important aspect of musical performance for many reasons, including synchronicity, expressiveness, and formal structure among others. Large-scale development is one aspect which the MAs do not operate well on their own. One strategy we have used is to change corpora in performances to give a sense of larger sectionalization. This

[172] Borgo, *Sync or Swarm, Revised*, 27.

implies the role of a computer technician or laptop performer (in this case, Cusson or Pocius) who changes the corpora during performances.

Regarding psychologist Jeff Pressing's research on improvisation, David Borgo states that "improvisers, according to Pressing, need only find a good solution, not the best, since the search for an optimum would be too time-consuming and resource intensive. Pressing, however, views this 'non-optimum' situation as providing the potential for unique outputs and novel interactions between musicians."[173] Derek Bailey notes that improvisers spend "very little time looking for 'new' things to play. The instinctive choice as well as the calculated choice is usually for tried material. Improvisation is hardly ever deliberately experimental."[174] What then is good enough regarding an MA's musical output while performing with an improviser? Above I have stated that the MAs' abilities pale in comparison to humans in certain ways, however, they may also extend beyond human capabilities in others. For instance, by combining various phrases together, they might output material that could not physically be performed by a human. The MAs are not bound by temporality like human improvisers. For example, their training phase does not take place in real-time. The MAs could also output the data for multiple different performances to a pre-recorded input. However, to listen back to these various performances, a human would listen in real-time and thus be bound by temporal restrictions. When one evaluates the MAs during or after a performance from an artistic perspective, there will inevitably be a comparison to past improvisations and interactions with humans as a kind of reference for how one interacts with the MAs.

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."[175] 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

[173] Borgo, 25.

[174] Bailey, *Improvisation*, 73.

[175] Ingrid Monson, *Saying Something: Jazz Improvisation and Interaction* (Chicago: University of Chicago Press, 1996), 26, http://www.SLQ.eblib.com.au/patron/FullRecord.aspx?p=432268.

Monson above as we must ask if computers have personalities. Paul Dourish states that interaction between a computer system and end-user is a "fundamentally a social activity" since the computer "mediate[s] communication between the end-user and the system designer."[176] 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.[177] 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

[176] Dourish, *Foundations of Embodied Interaction*, 56.

[177] Tatar and Pasquier, "MASOM."

embedded in the audio descriptor analysis. Ethical concerns related to biases inherent in these technologies are beyond the scope of this paper.

Our research into MAs involves an interdisciplinary team taking on different roles such as programmer, instrument builder, laptop performer, improviser, and composer, among others. As suggested above, these 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. As stated in Chapter 1, there is no optimal solution in Music Metacreation (MuMe), just like there is no optimal improvisation. However, thinking about what an optimal solution might be in a given situation, may shed light on what one finds important in improvisation or interaction. This would presumably also change based on the musical situation, other collaborators, and of course the technologies involved, and there may be many optimal solutions for any given situation. Considering possible optimal solutions may also be useful—what would a given situation mean for an improviser, what would one need to *do* to create an optimal solution, or what kind of behaviour would affect this kind of change in the music, for example. The optimal in this case depends on what we want to do artistically and is affected by many different variables that are often difficult to account for.

When performing with the MAs I would like to state that no one is in the lead, that I treat the agents as I would an improvising partner, and that we create the music together equally. However, as suggested above, this relationship is more nuanced and there exist layers of technological and human collaborator agencies that need to be peeled back.

When I say that no one is in the lead, this seems to contradict our approach to communicating with the agents via the controller, as the controller sends the MA’s commands which have a direct effect on the MAs. The MAs are capable of a certain type of interaction and music, and are autonomous in the sense that, once the programs are launched, they will continue to operate. Indeed, I do give direct commands via the controller, and take certain liberties over the MAs that I would not in a collaboration with other humans. I will address some of these issues throughout Chapter 6, culminating in Section 6.9 where I will introduce some ethical considerations when performing with MAs.

6.3 The Posthuman, Cyborg, and Media

There are many conceptions of the posthuman, and many with their own dedicated bodies of research that continue these original threads. A thorough engagement with even a small fraction of this literature is beyond the scope of this document, but I hope to show how the posthuman concepts have been useful for contextualizing improvisation with MAs as an initial step presented in this chapter. N. Katherine Hayles describes how "the posthuman configures human being so it can be seamlessly articulated with intelligent machines. In the posthuman, there are no essential differences or absolute demarcations between bodily existence and computer simulation, cybernetic mechanism and biological organism, robot teleology and human goals."[178] As suggested by Hayles, the posthuman does not refer to distinctly different organisms, although in some cases this may be the case, but rather it is a transition from the liberal humanist towards a more situated perspective. In this regard, Hayles states that "[the posthuman view thinks of the body as the original prosthesis we all learn to manipulate, so that extending...the body with other prostheses becomes a continuation of a process that began before we were born."[179] As suggested here, we are born at a specific time, including cultural and political factors, which determine what types of technologies are in use. Hayles states that "I understand human and posthuman to be historically specific constructions that emerge from different configurations of embodiment, technology, and culture."[180] Hayles' notion of moving away from human-centredness and focusing on how humans have become intertwined with technologies is relevant to working with MAs. In addition, the idea that these technologies may extend our bodies beyond the boundaries demarcated by our physical bodies provides a way to conceive of these interrelated social and inter-personal complications that come along with working with MAs. However, recent developments in interactive technologies using AI tools and human cognition and behaviour models, as with the machine listening discussion broached in Section 3.3.5

---

[178] N. Katherine Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics* (Chicago: The University of Chicago Press, 1999), 3, http://hdl.handle.net/2027/heb.05711.

[179] Hayles, 3.

[180] Hayles, 33.

regarding machine listening, may present outcomes or ramifications of these technologies that we do not yet understand. It is certainly exciting to be in the process of learning about these MA systems as they are also being developed and improved. The conception and construction of the posthuman vary widely, and the discussion presented below contributes one iteration of the posthuman and the cyborg as one posthuman paradigm. As stated by Hayles, "what the topology will reveal is not so much an answer to the deep question of how the human and the posthuman should be articulated together as the complexity of the contexts within which that question is being posed."[181] As with my discussion on improvisation, the current project is one example of how the MAs may be described as extending our human agencies as artists and as a collaborative team.

How does Hayles' formulation above help us to recontextualize and reconsider how we performer in human-human situations? How might we better understand the roles that we play in artistic collaborations following collaborations with MAs. As George E. Lewis articulates about his MA Voyager, "this work… deals with the nature of music, and in particular, the processes by which improvising musicians produce it. These questions can encompass not only technological or music-theoretical interests but philosophical, political, cultural and social concerns as well."[182] Lewis is the creator of Voyager, and so his perspective on this creation process is distinctly different than our process as adopters, users, and adapters of these technologies. However, the sentiment is noted, that we are indeed posing questions about music and the process of creating music, including considerations of what we might look for in human collaborators when improvising, while remaining open to distinctly machine musical propositions as well. Marshall McLuhan posits that technologies push the "archetypal forms of the unconscious out into social consciousness." This description by McLuhan continues to be relevant to the working process with MAs and our situation. As others have noted, I would also specify that this is a shared social consciousness, in which we are at first sharing internally within the team, as a way to decide upon a shared direction, or at the very least, share our own versions of an imagined artistic direction for the MAs. Second, we then share this group vision and MA agency with a larger

---

[181] Hayles, 251.

[182] Lewis, "Too Many Notes," 33.

public through performance. Depending on the performance context, we might also share specific details of our philosophical approach and the artistic process through a presentation, program note, or spoken introduction.

The term "cyborg" was first coined by Manfred Clynes in 1960 where he states that "the Cyborg deliberately incorporates exogenous components extending the self-regulatory control function of the organism in order to adapt it to new environments."[183] The word cyborg is a portmanteau combining the words cybernetic and organism, and in the cited article, Clynes discusses existing and hypothetical technological adaptations for human space travel. The cyborg metaphor has been explored in Donna Haraway's seminal Cyborg Manifesto where we are reminded the ways technologies are integrated into our lives and blur human bodily boundaries.[184] In a similar vein as Haraway, Marshall McLuhan discusses how one retroactively contextualizes and comes to understand older technologies through the process of adopting new technologies, since the "content" of a technology is always another technology.[185] These concepts from Haraway and McLuhan and their broad societal usage might be summed up by Lawrence Lessig's notion of "remix," where he describes the broadly accessible technologies that permit one to engage with create acts through editing, mashing, collage, and remixing of digital data into new artifacts.[186] The processes by which MAs reassemble and output audio, more specifically known as concatenative sound synthesis (CSS), interacts with recordings and affects our understanding of, and relationship to, recording technologies.[187] In the Eurological tradition, recordings have played an important role in establishing canonized works and interpretations which inform

[183] Manfred E. Clynes and Nathan S. Kline, "Cyborgs and Space," *Astronautics* 5, no. 9 (1960): 27.

[184] Donna Haraway, "A Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s," *Australian Feminist Studies* 2, no. 4 (March 1, 1987): 1–42, https://doi.org/10.1080/08164649.1987.9961538.

[185] Marshall McLuhan, *Understanding Media: The Extensions of Man* (1964; repr., London: Routledge, 2001), 8–9.

[186] Lawrence Lessig, *Remix: Making Art and Commerce Thrive in the Hybrid Economy*, 1st ed. (London: Bloomsbury Academic, 2008), https://doi.org/10.5040/9781849662505.

[187] Schwarz, "Concatenative Sound Synthesis."

ongoing performance practices throughout the 20th century. As Lawrence Kramer observes "one of the paramount features of 'posthuman' interfaces is the unprecedented volatility they offer by means of cutting and pasting, processes that…make texts, images and sounds easily transferable and transportable."[188] To use Lessig's terminology, the MAs remix the analyzed corpora, although with limited means, and this is the foundation for their sonic outputs. There is incredible flexibility with which we as creators can work with these materials via the analysis and interactive settings in software, including using records of pieces or improvisations across the spectrum of style and aesthetic, which is why we have set certain limitations related to source material for corpora and the eTube itself, as described especially in Sections 3.1.1 and 3.3.2.

Haraway calls attention to the scale of integrated machines in our society and states that "modern machines are quintessentially microelectronic devices: they are everywhere and they are invisible."[189] Haraway is correct in stating "invisible" as opposed to the narrative of digitalization being equated with the immaterial, as Kyle Devine shows in research on the ecological impact of recording technologies and digital media.[190] It was important for Cusson and I to design a wireless controller, to limit wires hanging from the eTube and trailing across the stage, to make allow the interaction with the MAs via an integrated controller. At times I do emphasize the gesture and key sound associated with eTube controller inputs, however the commands are not used constantly as the saxophone's keys are when playing. However, we anticipated the wireless technology would help a suspension of disbelief on behalf of the audience, and as part of our mise-en-scène, to contribute to the illusion that I appear to be interacting or communicating with the agents in performance.

In Section 3.3.5, I discussed how machine listening is a metaphor for human listening, and how there is no agreed upon definition or model for machine listening. As the Haraway example

---

[188] Lawrence Kramer, "Philosophizing Musically: Reconsidering Music and Ideas," *Journal of the Royal Musical Association* 139, no. 2 (2014): 392.

[189] Haraway, "A Manifesto for Cyborgs," 6.

[190] Kyle Devine, *Decomposed: The Political Ecology of Music* (Cambridge, MA: MIT Press, 2019), https://doi.org/10.7551/mitpress/10692.003.0001.

above forewarns, the technologies that we use are imbedded with biases. The interactive technologies we use, notably Max/MSP and the associated objects, is owned and managed by the company Cycling ’74, but certain aspects are also developed by an international community of developers.[191] OMax had been developed in Max/MSP at the *Institut de recherche et coordination acoustique/musique* (IRCAM), and is one of the foundational MA programs which would eventually lead to the development of DYCI2.[192] It is notable that the primary tool used in OMax for analysing a sound onset for a live input is a pitch detector and the separation of playing modes as “free mode,” with no metrical structure, and “beat-mode,” which references a metric and harmonic grid.[193] To a certain extent, the OMax successors SOMax, ImproteK, and DYCI2 are built upon similar assumptions regarding segmentation and separation of metric versus non-metric music, although these more recent MAs have significant updates in terms of timbral descriptors, for example. As I have stated above, I have made compelling and artistically fulfilling music with these MAs. Reflecting on my own practice, and others that I am familiar with, improvised music does not always lack pulse or rhythmic structure. In addition, pitch and harmonic material might not always be the most salient features for analyzing improvised music. An analysis or in-depth discussion of these questions is beyond the scope of this document. However, I do wonder how these assumptions surrounding metric vs. non-metric and a focus on pitch detection as a primary tool for segmentation might carry through from OMAX to ensuing versions of the software stated above, which may affect not only the type of music that is made, but also the type of music that is made possible with these MAs.

[191] For an in-depth ethnographic study of IRCAM, see: Georgina Born, *Rationalizing Culture: IRCAM, Boulez, and the Institutionalization of the Musical Avant-Garde* (Berkeley: University of California Press, 1995).

[192] For an in-depth ethnographic study of IRCAM, see: Born.

[193] Gérard Assayag, Georges Bloch, and Marc Chemillier, “Omax-Ofon,” in *Sound and Music Computing*, 2006, https://hal.science/hal-01161346.

6.4 The Intentional Stance

Daniel Dennett describes the concept of the intentional stance, wherein one agrees to understand an object's behaviour in terms of its outcomes, rather than understanding the physical properties or inner workings of the object.[194] Dennett describes three strategies for predicting future behaviour: the physical stance, the design stance, and the intentional stance. The physical stance describes how we can predict future behaviour based on the physical laws and properties of an object. For example, one can predict what will happen to a pot of stew if it is left on the stove at a high temperature. The design stance is a strategy where one assumes that a system will behave as it is designed to behave. For example, one can understand what a computer is designed to do and that it will function as designed, although we do not know how it actually works. The intentional stance requires that one treat an object as a rational agent and to pretend that the agent has beliefs and desires which serve its purpose. Based on these assumptions, one predicts how this agent will act to fulfill its beliefs, and this is what one predicts the agent will do.[195] The intentional stance attributes agency and intent so that we may understand the resultant outcomes as that of rational yet artificial agents. The intentional stance does not imply intentionality on the part of the agent, but that "the user might be able to intuitively and reliable [sic] explain and predict their behavior in these terms."[196] The intentional stance is a heuristic that allows humans to rapidly understand specific situations, although it does not provide the same detailed information as the physical or design stance.[197]

---

[194] Daniel C. Dennett, *The Intentional Stance* (Cambridge, MA: MIT Press, 1987), 13–42, https://proxy.library.mcgill.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=48751&scope=site.

[195] Dennett, 17.

[196] Perez-Osorio and Wykowska, "Adopting the Intentional Stance toward Natural and Artificial Agents," *Philosophical Psychology* 33, no. 3 (2020): 372, https://doi.org/10.1080/09515089.2019.1688778.

[197] Perez-Osorio and Wykowska, 384.

Applying the intentional stance to MAs as Eric Lewis suggests, one pretends that MAs have beliefs and desires, and that they are indeed improvising, and to engage in improvisation as one would with another human improviser. As stated by Marc Leman, interactive technologies as an extension of the human may increase peak experiences where users are "totally immersed in energy."[198] An important aspect of interactive experiences is for the software to generate the right content at the correct time for the user. Flow states are described by Mihaly Csikszentmihalyi which are optimal experiences where the user or performer is immersed in the given activity, allowing for deep focus and involvement while flow is maintained.[199] Users will require different interactions with the software in order to maintain flow states and this is determined by their previous experiences with the systems, and in the case of improvisation with MAs, their musical experience. One of the justifications for real-time generation is to induce flow states in users through the creation of appropriate content to maintain flow.

6.5 Make-Believe and Fiction

Using the intentional stance, we pretend that an artificial agent has desires and beliefs and use these to predict its behaviour. The crucial distinction made by Lewis is that we need to consider the difference between believing *as if*, and make-believe, and for this Lewis suggests we consider how one relates to fictional characters.[200] One must then ask, what constitutes fiction in the first place? Gregory Currie states that "it is not any linguistic or semantic feature of the text that determines its fictionality" and one does not simply pretend to assert a fictional discourse.[201] The difference between fiction and non-fiction according to Currie, is that the author intends "we

198 Marc Leman, *Embodied Music Cognition and Mediation Technology* (Cambridge, MA: MIT Press, 2008), 140, https://doi.org/10.7551/mitpress/7476.001.0001.

199 Mihaly Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*, First edition (New York: Harper & Row, 1990).

200 Lewis, *Intents and Purposes*, 97.

201 Gregory Currie, *The Nature of Fiction* (Cambridge: Cambridge University Press, 1990), 12, https://doi.org/10.1017/CBO9780511897498.

make-believe that the story as uttered is true."[202] Following Currie's statement above, in our improvisation context, I do not *pretend* to improvise with the MAs onstage, I am indeed improvising and I imagine that the MAs are indeed improvising with me!

As stated in Section 1.3 above, positions such as John Searle's Chinese Room argument pose a challenge to claims of strong AI, described as machine intelligence that is indistinguishable from the human mind. The MAs operate in compelling ways during improvisation, but to say that they *actually* improvise would seem to involve some kind of intent on the MA's behalf. In most ways the MA's abilities pale in comparison to a human improviser. Searle's argument complicates the issues of whether a machine can "understand" or exhibit "intention." As is suggested by Eric Lewis' account above, it is the human's role to imagine the "intent" that an MA displays in an improvised performance.

Humans engage regularly with fictional characters, and as proposed by Lewis, we may indeed consider MAs as imaginary partners and "make believe we are improvising with them with little if any loss of authenticity."[203] It is also due to the success of the technologies we employ that we are able to engage in this make-believe. Is it then a question of complexity that helps define when one is improvising? If one used complexity as a benchmark for improvisation, how are we to explain artists that improvise with sine wave drones produced by analog synthesizers, a technology originally designed for military use, and comparatively less complex than musical agents. Are we to say that these artists are not improvising because they are producing less complex sounds relatively speaking? Todd Winkler states that a free improvisation MA should "creat[e] listenable music on its own."[204] Should a subjective judgement of the music being able to stand on its own be necessary for improvised music? If an improviser hated the music that others were performing, a bold statement might be to sit on stage and to play nothing. This statement would not be related to a musical output, or any subjective notion of the music being

[202] Currie, 18.

[203] Lewis, *Intents and Purposes*, 98.

[204] Winkler, *Composing Interactive Music*, 26.

able to stand on its own but would remain a bold musical statement. One might claim that processing power and real-time processing allows us to consider that the MAs are improvising. But what is the difference between computer assisted composition and the MAs? Prior to today's processing power, we might have input audio into similar software, then let it process for a week, before coming back to listen to the output. Would we still consider this situation as improvisation, or would one be more likely to consider it as computer assisted composition? One of the main differences between the above situation, and performing in real-time with MAs, is that the temporal aspect is shortened by an increase in processing power available to MAs today. However, is the fact that these computational processes occur almost instantaneously sufficient to suggest that a computer is actually improvising?[205]

Many people in today's world traverse in and out of various virtual spaces daily. This includes engaging with so-called "smart" devices or virtual profiles on social media. The magic circle is a concept from game theory which is defined by Joshua Fairfield as "the supposed metaphorical line between the fantasy realms of virtual worlds and what we consider to be the real world."[206] These virtual boundaries allow users to experience play within the game's bounds and involve following rules which are often culturally determined.[207] Recently, virtual spaces are also shared during concerts performed by holograms of deceased artists. Based on hologram concerts in Japan and the US, Yuji Sone shows that audiences' reactions to virtual concert spaces are culturally influenced.[208] These spaces are ubiquitous in much of today's world, and one may take part in various virtual spaces, transitioning between gameplay and cooking dinner, for example. If we are to make believe that the agents are also improvising with us, we are also implying that the audience might do the same. In our project, we are inviting the audience to share in these

---

[205] Thank you to Eric Lewis' insight for these examples.

[206] Joshua A. T. Fairfield, "The Magic Circle," *Vand. J. Ent. & Tech. L.* 11 (2009): 824.

[207] Jaakko Stenros, "In Defence of a Magic Circle: The Social, Mental and Cultural Boundaries of Play," *Transactions of the Digital Games Research Association* 1, no. 2 (2014).

[208] Yuji Sone, "Dead Stars and 'Live' Singers: Posthumous 'Holographic' Performances in the US and Japan," in *Sound and Robotics*, ed. Richard Savery (Chapman and Hall/CRC, 2022), 317–36, https://doi.org/10.1201/9781003320470.

tacit agreements about the MAs in the concert space. As discussed above, we are used to interacting with virtual personae via our mobile devices and online avatars on social media stand as representations of ourselves in an online environment. To ask the audience to pretend that the MAs are improvising with us, although we understand this situation to be a fiction, is not so unfamiliar in this age of tech and media—to ask audiences to pretend is to invite them to join our game, or to play along in a shared virtual space.

Peter Lamarque and Stein Haugom Olsen's "internal" perspective relates to fictional characters as having the range of properties of real people (e.g., being arrogant, wily), versus the "external" point of view where the properties belong to a different category (e.g., being created by an author).[209] The internal context relates to the discussion above, where I make believe the MAs have beliefs and desires, like a human. The "external" perspective would refer to the MAs as a technology which is constructed, shaped, and contextualized by human programmers and artists. As an artist working with MAs, I am also engaged in this dance between internal and external. At times I work to artistically evaluate the MAs and how I interact with them in performance, including how they have been designed by the original programmers. In addition, I also engage with them in performance, making believe that they are interacting and improvising as a human might. Flint Schier adds that "we are reacting to characters as vividly seen and realized by a controlling intelligence and we respond to the work as an expression of that achieved vision of the characters."[210] From the audience's perspective, their comprehension of the MAs as would be based on their own previous knowledge and experience with these technologies and influenced by how we present them through spoken text, program notes, and in the performance (see Section 6.9).

If we decide to make-believe that MAs do indeed improvise with us, then we still need to address the fact that the MAs are disembodiment, as described above. Theory of Mind (ToM) research

[209] Peter Lamarque and Stein Haugom Olsen, *Truth, Fiction, and Literature: A Philosophical Perspective* (Oxford University Press, 1996), 146, https://doi.org/10.1093/acprof:oso/9780198236818.001.0001.

[210] Flint Schier, "Tragedy and the Community of Sentiment," in *Philosophy and Fiction*, ed. Peter Lamarque (Aberdeen: Aberdeen University Press, 1983), 85–86.

investigates how we come to know people and learn about their emotions, beliefs, and intentions, which is crucial to human social interaction,[211] and it is well recognized that this happens through facial recognition.[212] ToM researchers are interested in how quickly we can read emotional states based on facial recognition. Gunnar Schmidtmann and colleagues show that humans can reliably associate mental states from facial expressions showing the eye region only, even when the facial expression is visible for only a fraction of a second.[213] What would a ToM for MAs consist of? As artists working with MAs, what would we want to know about them and by what senses would we want to be informed? The aforementioned ToM research shows that these cognitive recognition processes are intricately linked to human physiology—viewing another person's eye region. It seems that if one was to perform some equivalent of ToM with MAs, it would require a certain amount of experience with the MAs. Andrew Brown and colleagues suggest that one might develop "algorithmic listening," or the ability to hear the "*procedure* causing a musical event" (italics in original).[214] Anecdotally, the authors have experienced something that might be considered algorithmic listening. While working with the MAs we have identified certain behaviors (or lack thereof) where it was clear the specific function in the patch that was working (or not working).[215] It is perhaps because of our experience with the agents through extensive use in performances and rehearsals, that we may be able to develop an intuition or deeper understanding about the MA's functioning and output. As suggested by Bown's algorithmic listening, this new understanding might be likened to an embryonic ToM for MAs on some level.

---

[211] Lisa Zunshine, "Why We Read Fiction," *Skeptical Inquirer* 30, no. 6 (2006): 29.

[212] Gunnar Schmidtmann et al., "In the Blink of an Eye: Reading Mental States From Briefly Presented Eye Regions," *I-Perception* 11, no. 5 (2020), https://doi.org/10.1177/2041669520961116.

[213] Schmidtmann et al.

[214] Brown et al., "Interacting with Musebots," 22.

[215] Davis et al., "Case Studies."

6.6 Embodying Agents Through Spatialization

Eric Lewis summarizes Jerrold Levinson's "musical personae" theory[216] by describing how "music's expressiveness is a product of a listener imagining that the music is literally expressing emotion."[217] Eric Lewis continues his argument on the intentional stance and fictional characters when he states that "*all* music listening requires one to take such an imaginative leap, to hear an imaginary persona in the music."[218] If all musical listening involves imagining a persona, then audiences and performers alike take part in a shared, yet subjective listening fiction. For Paul Sanden, this is the "productive tension" between "traces of live performance" and mediatization which allows one to find meaning.[219] We can imagine that the audience would engage in some way with both biological and artificial performers' personae. But what might they comprehend about the MAs and technical details like the corpora and the ways I interact with the mic? As suggested by Sanden, audiences might hear the traces of bodily resonance, inhalation, or the singing voice in corpora segments output by the agents. Although the MAs are highly mediatized and disembodied, although perhaps embodied simply as the user interface on the computer screen, we have intentionally left these traces of my corporeality in the corpora, and as a result, they may be heard in the MA's output. As stated by Flint Schier above, audiences learn about fictional characters through an engagement with the controlling intelligence, in our situation, this would mean engaging with the onstage performer. However, audiences also learn from spoken introductions or program notes which might communicate notions such as fictional characters, agency, and gestures surrounding the microphone, in addition to the aspects of the actual performance and artistic presentation. Each audience member will interpret these details through their own subjective experience depending on their understanding of the technologies,

[216] Jerrold Levinson, "Musical Expressiveness as Hearability-as-Expression," in *Contemporary Debates in Aesthetics and the Philosophy of Art*, ed. Matthew Kieran (Malden, MA: Blackwell, 2006), 192–206.

[217] Lewis, *Intents and Purposes*, 98–99.

[218] Lewis, 98.

[219] Paul Sanden, *Liveness in Modern Music: Musicians, Technology, and the Perception of Performance* (New York, NY: Routledge, 2017), 113.

performance techniques, and music, which will affect how they experience the roles and interaction during the performance.

As discussed in Section 4.1, Pocius' spatialization systems use pitch and amplitude information extracted from the mic signal to control the localization and movement of the agents throughout the performance. This allows the MAs to adopt movements throughout the loudspeaker system in relation to acoustic performers' playing style without the need for additional sensors, with the possibility to map the data to the movements of each agent differently helping to keep them separate for an audience member. The impetus for creating these models was to engage with my spatialized performance practice, which was Pocius' initial inspiration when they created their models. Michel Chion states that "a large majority of visible things remain constant, whereas a large majority of audible things are temporary."[220] However, although audible things are temporary, when a sound source moves, it is much more easily localized compared to a static sound precisely because its location is always changing.[221] Pocius has organized the spatialization such that the effects processing from Construction III is placed between the loudspeakers, and the MA's outputs are placed closer to the loudspeakers. The movement from the MAs then, is tied more strongly to the physical speakers, anchoring the MA's sounds in the hall's features. Spatialization is one way to interact with the virtual space of the concert hall, and by moving sound through the hall as a cohesive unit with the models, may create the sense that the agents are embodied by the hall's structures, or that the sounds are moving as a cohesive unit.

6.7 Interacting with the Microphone

The microphone could be equated in a simple metaphor as the "ear" of DYCI2 agents since this is how they analyze the audio signal from the eTube. Put another way, the microphone enables the MA's "umwelt." The word umwelt was coined by Jacob von Uexküll and refers to how each

---

[220] Michel Chion, *Sound: An Acoulogical Treatise*, trans. James A. Steintrager (Durham: Duke University Press, 2015), 37, https://doi.org/10.1515/9780822374824.

[221] Chion, 25.

biological organism maps and understands its own specific way of perceiving the world.[222] This notion has been extended to a "technological umwelt," as suggested by Rosemary Lee, to refer to the perceptual apparatus that a non-biological agent would use to take in information from its environment.[223] In our project, this apparatus is the microphone. To maintain consistency, my standard setup uses an Electro-Voice RE20 mic on a stand which allows me to influence the agent's umwelt by moving the eTube around the mic while performing. However, we have tested other microphones with the eTube. For example, I have used a clip-on mic attached to the sounding end of the eTube. A clip-on mic maintains a direct and consistent sound input to the mic regardless of the eTube's position or orientation. With both mic setups, I may still influence the agents through dynamic fluctuations and varied performance gestures. However, with a clip-on mic, the MA's umwelt stays relatively stable, due to the consistent input to the mic. As mentioned in Section 3.2.1, each DYCI2 agent has an adjustable amplitude threshold that modifies the input necessary to provoke an output from the agent. When performing with the clip-on mic, the DYCI2 agents would activate too easily and even the softest eTube sounds would launch an output. Therefore, clip-on mics did not allow me the possibility to perform solo material without MA interjections.

In contrast to clip-on mics, with the RE20 on a stand, my spatialized performance gestures have a distinct effect on the MA's umwelt, since the mic's input will vary depending on where the eTube is oriented in the space. I had performed certain physical gestures with tubes before the eTube project, but the stationary microphone has influenced the addition of physical gestures specifically for the eTube mic setup and considering the artistic affordances of the MA's

[222] Jakob von Uexküll, "A Stroll through the Worlds of Animals and Men: A Picture Book of Invisible Worlds," *Semiotica* 89, no. 4 (1992): 319–91, https://doi.org/10.1515/semi.1992.89.4.319.

[223] Rosemary Lee, "The Limits of Algorithmic Perception: Technological Umwelt," in *Politics of the Machines-Art and After* (Aalborg University, Copenhagen, Denmark: BCS Learning & Development, 2018), 2, https://doi.org/10.14236/ewic/EVAC18.44.

threshold values.[224] From my experience performing with the saxophone and a variety of microphones, I would intuitively move the eTube closer to the mic when I wanted to activate a DYCI2 output. When playing louder material, if I did not want DYCI2 to respond, I would direct the eTube away from the microphone. This approach allowed me to shape the MA's output independently of my sonic output by directing the instrument in space relative to the mic.

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 I 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.[225] If I purposefully perform gestures away from the microphone, I 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 I have spoken about the MAs being co-collaborators, this is not always an equal collaboration (see Section 6.9), and I chooses to assert this control over the MA's umwelt as an artistic affordance.

Human performers use subtle gestures and visual cues to communicate during performances. As stated above, the MAs we use do not sense visual data, they only sense the performance environment via the mic. The MA's umwelt comprises fewer modalities than a human, presenting limitations in terms of how I can interact with the MAs, compared with a human improviser. However, Arne Eigenfeldt and Oliver Bown's Musebots demonstrate how MAs may communicate with each other during a performance to synchronize tempo, attacks, pitch sets, or density, for example.[226] Thus, a Musebot ensemble may negotiate change and communicate in

---

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

[225] Lee, "The Limits of Algorithmic Perception: Technological Umwelt," 1.

[226] Arne Eigenfeldt et al., "Distributed Musical Decision-Making in an Ensemble of Musebots: Dramatic Changes and Endings," in *Proceedings of the International Conference on Computational Creativity*, 2017, 88–95.

their own digital ecosystem, going beyond human communication capabilities. An improviser would not, in general, place their instrument directly on another improviser's face or ears in the way I place the eTube directly on the microphone grill. However, whereas I can communicate with another improviser using subtle gestures, broader gestures often seem more appropriate when interacting with the MAs. This seems to be a difference of degree and not of kind, which is influenced by the MA's umwelt and apparatus—the microphone. I take the liberty to use these performance gestures to influence and control certain aspects of the MA's performance.

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, and later discusses how a contact mic would have been better than a "normal microphone" to record the piano input for Spire Muse.[227] 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.[228] Based on this documentation, the developers refer to a consistent input, with no mention of interacting with the MAs via the microphone as I have described above.

The notion of the black box, where the input and output are clear, but the internal processes resulting in the output are not evident, has been introduced in Section 3.3.5. I consider the space around the microphone that I interact with as a metaphorical black box space. Although I know there are threshold values, I do not know exactly at what amplitude at a specific distance from the mic will launch a MA output. I keep the mic setup as consistent as possible, and use an RE20 when possible, however, other variables such as the room acoustics and loudspeaker placements also affect the MA's threshold. I learn at what dynamic, and at what proximity to the mic each of the MA's thresholds are toggled as I perform with the MAs in real-time. I engage with this metaphorical black box, or conceptual threshold space, by adjusting my performance gestures and movement in relation to the kinds of interaction I want. However, I question how audiences understand the interaction between the improviser and mic, and if it the relationship between mic

[227] Thelle, "Mixed-Initiative Music Making," 180 and 234.

[228] Nika et al., "Dicy2 for Max," 12.

input levels and MA outputs are apparent to the public. Mediatized concerts are the norm today, and audiences are used to performers interacting with microphones, adjusting their bodies, or moving the mic for phrasing, dynamic, and theatrical effect. How the audience understand the MA as a black box is subjective based on their technological and musical knowledge and experience, which may be different from how members of the eTube team understand these elements. As a result, I often emphasize certain movements around the microphone to accentuate triggering the MA's threshold and launching a response. For example, I slowly move the eTube towards the mic, and as soon as I perform a slap tongue attack, I quickly pull the eTube away from the mic. This may give a rhythmic emphasis to these repeated gestures that are marked by the percussive slap tongue attack. If the MA performs something that I would like to respond to immediately, I must quickly move the eTube close to the mic. I may pull the instrument away from the mic to allow the MA to respond, before quickly placing the eTube in front of the mic to respond back. These ancillary gestures are not directly related to sound production but are movements that interact with the mic, which also might communicate to audiences when I intentionally launch an MA's sonic output.[229]

6.8 The Conceptual Nod

Based on past improvised performances with saxophonist and media artist Nick Zoulek as the Duo d'Entre-Deux, I have been thinking about how certain types of visual feedback between performers could be adapted for use with MAs. Since 2011, the Duo d'Entre-Deux has developed a unique voice, improvising together for collaborative creation projects and in live concerts. We have co-improvised works together, including site-specific dance performances with Wild Space Dance mentioned in Section 2.1. Recently, we co-improvised a piece for two tenor saxophones with Montreal-based composer and violinist Alissa Cheung.[230] Throughout

[229] Marcelo M. Wanderley et al., "The Musical Significance of Clarinetists' Ancillary Gestures: An Exploration of the Field," *Journal of New Music Research* 34, no. 1 (2005): 97–113.

[230] "Alissa Cheung, Violin," Alissa Cheung, Violin, accessed March 28, 2024, http://www.alissacheung.com/.

these projects we have spent hours improvising together. We know each other very well, and like many improvising partners, we often understand where each of us wants to go musically, from sonic indications alone. However, body language and gesture remain an important communication mode in performance. In my experience performing with Zoulek, nodding the head has been a significant gesture which I often interpret as, “yeah,” “keep developing this idea,” or “you’ve got this.” In general, a head nod indicates that we are on the same track, or that Zoulek is encouraging the direction that I am taking, and he expresses his approval and that I should continue in a similar direction. This is indeed a physical gesture that may be observed in many types of performance situations in general, and outside of performance.

Thinking about performing with MAs, how would an agent give this type of feedback to an improviser? Firstly, on what metrics would this kind of feedback be based, since as we have discussed previously, the agents have a limited “understanding” of music or musical structure and their analyses and output are primarily based on models of listening and audio descriptor analysis. Notto Thelle implemented a *Thumbs Up* feature in Spire Muse and CCCP which signals that the performer is enjoying the interaction and keeps the agent in the same state for the next 30 seconds.[231] Jon McCormack and colleagues studied a bi-directional communication system where real-time skin conduction biometrics was communicated to the AI improviser, and emojis on a screen were used to communicate the agent’s confidence level to the improviser.[232] These approaches attempt to replace the non-verbal communication between humans with various sensors and processes designed to provide feedback and to ameliorate the interaction in performance. What would it look like for a MA to give a nod to an improviser and vice versa? I have named this idea the “conceptual nod,” a proposed structure to allow the MA to communicate with the improviser. This could also be implemented as controller input allowing me to also “nod” to the MA, indicating to the agent to “follow that way,” to borrow Jérôme Nika’s phrase.[233] As for the MA, perhaps its feedback could also be related to a statistical

---

[231] Thelle and Pasquier, “Spire Muse.”

[232] Jon McCormack et al., “In a Silent Way: Communication Between AI and Improvising Musicians Beyond Sound,” 2019, 1–11, https://doi.org/10.1145/3290605.3300268.

[233] Nika et al., “DYCI2 Agents.”

analysis of its corpora output over the course of an improvisation. An MA might be programmed to communicate that certain corpora segments are not being used and that a change of material is in order, a kind of nudge for the improviser to try something surprising. This could be programmed as a pulse that ramps up in intensity via a tactile sensor on the eTube, for example. I mention this here as future work for our project, and as a way to conceptualize how we might implement two-way communication and feedback between the improviser and the MAs.

6.9 Ethical Concerns and Care

Although we may be asking the public to share in these performative fictions with us, we feel it is nonetheless important to be clear about the MA's capabilities. Attribution is an important aspect of presenting the MAs to the public. For example, when we speak of the MAs as being trained on audio corpora, it may not be clear to audiences who performed on the corpora, who made the recordings, how and when corpus training occurs, and the MA(s) and human agents responsible for these tasks. Corpora creation includes improvising and recording sound for the corpus, curating, and editing the recorded audio, and determining the machine listening settings for MA analysis. As described in above sections, these tasks are often undertaken or shared by different team members, depending on the performance. Explicitly acknowledging these details with audiences or through written documentation highlights the human involvement throughout the process and credits the specific people responsible for these tasks. This is turn allows one to communicate which aspects of the performance are determined by the MA more clearly. We strive towards this goal, although I will be the first to admit that these concepts and processes are challenging to understand after dedicating three years of study, let alone after a single evening concert, program note, or presentation. However, this does not and should not prevent us from improving our communication so the concepts and delivery may be as clear as possible for the specific public we are addressing. Some may argue that by informing the audience about these details, that this would ruin the spectacle aspect. David Borgo agrees when he states that "media coverage of [artificial intelligence] also tends to promote sensationalist views that favor hype over reality, using dystopian or utopian rhetoric in place of real critical engagement."[234] Musical agents are distinctly different than AI tools such as ChatGPT, which are trained on immense language models. While a discussion of AI and the effect of this technology is beyond the scope of this paper, I do think that considering the ethical ramifications of these technologies is critical. Regarding the MA technologies we work with, rather than resting on dystopian narratives of technology, I prefer to present the MAs as our team treats them in this project, as collaborators which are both products and extensions of various human agencies.

We strive to learn about and respect how the MAs create music in performance which must be balanced against the impulse to always improve the MAs in what we view as musical deficits. Thelle and Wærstad describe the contradiction between designing an MA and wanting it to also be a co-creator in performance. We have the power to implement "'solutions' in the machine," whereas correcting other human performers is more "abstract" involving necessary "social codes and communication."[235] Although we are not the original developers of the software, we are nonetheless implementing new functions, updates, and determining how the MAs are used in performances. Part of considering the MAs as co-collaborators is taking the time to perform with and learn about them. In this respect, Oliver Bown states that "the participant in the interaction gains a direct sense of the interactive nature of the system, that may be obscure from outside."[236] Without this direct exposure to the MAs, we might miss out on key experiences and observations about the MA's behaviour, and the opportunity to feature these behaviours artistically. Avoiding the gut instinct to immediately fix what we might view as shortcomings or errors in the MA's behaviour allows one to consider how these limitations might be utilized or showcased in performance. As biological agents working with artificial agents, we have the power to make these changes. One would not stop a rehearsal with a colleague and insist on practicing long tones for hours to improve what one views as a deficit in their colleague's tuning. Although an equivalent situation might occur when working with MAs! This is of course from our perspective

234 Borgo, *Sync or Swarm, Revised*, 202.

235 Thelle and Wærstad, "Co-Creative Spaces."

236 Bown, "Player Responses to a Live Algorithm: Conceptualising Computational Creativity without Recourse to Human Comparisons?," 131.

as primarily users of these technologies, the developers would have had a much different process when creating the MAs. As users who update these technologies, although we could implement new code or modules to improve the MAs where we sense some deficit, I feel it is important to spend time with the MAs to reflect upon the changes that could be made, and the reasons for making them. Regardless, the original developers are continually improving and updating their software so the eTube team must none the less learn about these new updates and new affordances. In the midst of these continual updates and technological troubleshooting, it is a relief to simply play with the MAs at times, try new strategies, rehearse, and learn about the agents without feeling the need to ameliorate and improve the software.

Referring to MAs, Michael Young and Tim Blackwell state that "both human and machine contributors must have equal status."[237] What does equal status mean with MAs? First, one must assume that "equal" is only a consideration in specific contexts or environments. Having equal status with the MAs may seem to have an intuitive appeal, but there is a distinction between being equal in collaboration and equal in status. In Eurological musical creation, there are different hierarchies and statuses at play.[238] The role of the composer, compared to that of the performer, for example. In our project, we focus on corpora training since this contributes in substantive ways to how the MAs interact, for example. But how does the person who trains the corpus relate in status to the MA or the human improviser? While we aim for the MAs to be equal in status, in performances the MAs often follow the performers more than they are followed by the performers. The MAs will not begin outputting sound without an input, or by being toggled directly in the software, and they cannot end an improvisation on their own. In this respect, George E. Lewis describes "the customary practice of ending *Voyager* performances by

[237] Michael Young and Tim Blackwell, "Live Algorithms for Music: Can Computers Be Improvisers?," in *The Oxford Handbook of Critical Improvisation Studies, Volume 2*, ed. Benjamin Piekut and George E. Lewis (Oxford: Oxford University Press, 2016), 507–28, https://doi.org/10.1093/oxfordhb/9780199892921.013.002.

[238] Lydia Goehr, *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*, Revised edition. (Oxford: Oxford University Press, 2007), https://ebookcentral.proquest.com/lib/mcgill/detail.action?docID=3053152.

simply turning the machine off and having the humans create the endings."[239] In our project, the performer(s) also manage the endings in collaboration with the technician or laptop performer since the MAs are not capable of stopping a performance voluntarily, not to mention navigating an ending with other performers.

When I refer to the MAs being considered "equal" in this project, I mean the possibility of being considered equal in a specific environment, namely performances. As stated above, I use the intentional stance and make-believe to pretend *as if* the MAs are improvising with us. Although we are engaging with these MAs as if they are improvising, we are also projecting human abilities, such as intention, onto the agents in order to engage in a specific artistic scenario. The MA software do indeed have complex algorithms, models of musical emotion, and processes that inform their behaviour and sonic output. However, I want to be clear that the MAs are not equal to humans in most, if not all, circumstances outside this limited performance context. This is the distinction between being equal in collaboration versus being equal in status. If there was an emergency, there would be no question whether one would save the MAs, or a human life. I would not give a second thought to leaving the agents in a burning building to save my own, and other human lives. If I left my computer containing the MAs in a burning building, I might consider the financial loss, or the loss of precious information such as family photos, if the computer was destroyed. However, I would not consider that the MAs equal in status or value compared to a human.

Unlike in improvisations with humans, we cannot learn from the MAs through dialogue. As shown by Amandine Pras, improvisers may have distinctly different characterizations of specific moments in improvisations that they participated in.[240] Although certain shared musical understandings seem necessary for a performance to operate, Pras shows that improvisers often disagree on "statements about performers' thoughts and actions" as they reflect and recall

[239] Lewis, "Co-Creation: Early Steps and Future Prospects."

[240] Amandine Pras, Michael F. Schober, and Neta Spiro, "What about Their Performance Do Free Jazz Improvisers Agree upon? A Case Study," *Frontiers in Psychology* 8 (2017): 1–19, https://doi.org/10.3389/fpsyg.2017.00966.

specific moments in the improvisation.[241] That a human and MA's perspective on the performance is inevitably very different, and in many ways we are not able to communicate an agreed upon result with the MAs at all, but this non-agreement should not prohibit an improviser from forming their own opinions about a successful performance. While learning about the MAs, we may enter into some kind of dialogue with the original developers with the explicit understanding that the MAs have been designed for specific performance and aesthetics, they have inherent biases, and at this current time they do not yet demonstrate life-long learning without human intervention.[242] However, research on Artificial Life (A-Life) by Eduardo Miranda and colleagues shows how generative programs may operate in certain ways without human intervention.[243]

6.10 Case Study: Revisiting NIME 2022

I would like to return to a musical example, which after all, is the primary driving force behind my research and this project. This experience happened early in the process, and to which Cusson, Pocius, and I have often referred to as an important, meaningful, musical, and surprising cohesive experience performing with MAs.

---

[241] Pras, Schober, and Spiro, 17.

[242] George E. Lewis, "Why Do We Want Our Computers to Improvise?," in *The Oxford Handbook of Algorithmic Music*, ed. Roger T. Dean and Alex McLean, vol. 1 (Oxford University Press, 2018), https://doi.org/10.1093/oxfordhb/9780190226992.013.29; Lauren M. E. Goodlad, "Editor's Introduction: Humanities in the Loop," *Critical AI* 1, no. 1–2 (October 2023), https://doi.org/10.1215/2834703X-10734016; Jonathan Sterne and Elena Razlogova, "Machine Learning in Context, or Learning from LANDR: Artificial Intelligence and the Platformization of Music Mastering," *Social Media + Society* 5, no. 2 (2019), https://doi.org/10.1177/2056305119847525.

[243] Eduardo Reck Mirandau, *A-Life for Music: Music and Computer Models of Living Systems*, ed. Eduardo Reck Miranda, Computer Music and Digital Audio Series; Volume 24 (Middleton, Wisconsin: A-R Editions, 2011), https://ebookcentral.proquest.com/lib/mcgill/detail.action?docID=3115105.

There have been some notable experiences with the MAs which have been both shocking and exciting. In the following case, we were fortunate that this experience happened during a video recording session for our New Interfaces for Musical Expression (NIME) 2022 online performance.[244] Please refer to the timestamp (05:20–08:31) in the cited video link of our NIME 2022 conference performance. This video was made up of multiple takes of different improvisations during the day, and it is manufactured using a multiple camera setup in an attempt to give the impression of one continuous improvisation. However, this section noted above is part of one complete take, with no edits to the MA or my performance. I begin this short section by introducing slap tongues, which are responded to by slap tongue segments from the MA. This would be expected based on the MA's machine listening strategies. This is not in itself outstanding; this is what the MAs are designed to do. But what I do suggest below is that this improvisation created a meaningful musical and artistic experience for me, which seemed to be similar to a kind of breakthrough in the process of working and interacting with the MAs. This might be compared to the kind of experiences I have had as a performer and improviser when I finally find that sound I have been searching for, or I have broken through a barrier in some technique, musical idea, or interpretation. In this case, the breakthrough was both partially imposed by the MA, and perhaps also a limitation in my own understanding and experience, or perhaps a short-sightedness, in how I was working with the MA. When I first heard MA's rhythmic slap tongues in response to my own slaps, I committed to continue that one idea, in the same way that I would propose an idea in an improvisation with another human, by trusting the other person, proposing an idea, sticking with it, and developing it, while adjusting and reacting in real-time based on the musical propositions of the other improviser.

Especially significant is 07:18–07:45, where there is a notable section of interplay between repeated slaps, which upon listening, seems to be first initiated by the MA's slow slaps introduced at 07:18. I am always intrigued when listening to this section at the rhythmic interest and vitality between the two parts, but also how the two parts interact and weave through each other in a spontaneous and musical way. This is of course my own subjective opinion, it was a striking moment when I performed it, so perhaps it is coded in my memory as being meaningful,

---

[244] *eTu{d,b}e by Vincent Cusson and Tommy Davis - NIME 2022*, YouTube video, 14:12, 2022, https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.

which then affects my reaction to repeated listening. At this stage in the project, I began to be frustrated with the MAs for the seeming lack of ability to play rhythmically (this is perhaps also a reflection of my corpora recordings, see below) and in longer more cohesive phrases. I was so frustrated that in the middle of the recording session, I decided to record a new corpus on tenor saxophone (not heard here), which consisted of rhythmically consistent articulated and slap tongued notes as a strategy to increase the amount of rhythmic material available in the MA's corpora. As discussed in Section 3.3.2, this rhythmic tenor saxophone corpus documents the process of working with the MA, and the approach that I took as an improviser to encourage a certain kind of rhythmic interaction and audio output from the MA. However, this specific section from the NIME 2022 improvisation was a serendipitous moment as one of my main frustrations with the MAs—rhythmic interplay—was countered, and there seemed to be some kind of a breakthrough in the collaboration with the MAs. I am not suggesting any notion of intent on the part of the MAs, they are neither able to understand nor respond to my frustration with their lack of rhythmic variety in any concrete or communicative way. Rather, I am posing a more general question about the ways this collaborative process has unfolded, and the ways that considering the MAs as a co-collaborative partner allows me to offload certain responsibilities onto the MA, in a similar way that I would trust another improviser to be responsible during a performance, and the artistic affordances of this relationship.

As Borgo suggests, in referent-free improvisation, team-reasoning skills aid to create a kind of "scaffolding" upon which the free-improvisation may be constructed.[245] Part of the reason I attribute to this success is that I was intimately aware of the types of material in the MA's corpora for this performance. I was also familiar with the MA's behaviour from my experience working and playing with them. I recorded the corpora, and I knew that there were extensive sections including slap tonguing on baritone saxophone and eTube, and I had improvised with these in previous sessions. If the material of the corpus is my own and is of similar material that I would usually improvise with, this would also suggest that it should fit well with my own improvising. Although, this self-similarity between my playing and the MA's corpus is not the whole story, because as we know, the MAs operate in non-deterministic ways, and so the agent

[245] Borgo, *Sync or Swarm, Revised*, 180.

is also contributing to how the improvisation unfolds and develops. This is not in the same way that two humans would agree to rehearse a musical excerpt or to agree on a specific musical idea, however. I also attribute this success of this excerpt to the knowledge I have gained regarding how the MA's make music, and how that affects my own interaction and performance with the MA as a result. This is indeed part of the original conception behind the "etude" in the *eTu{d,b}e* framework, that the artists are constantly learning from the agents and learning about the specific types of music proposed or suggested by interactions with the MAs.

Derek Bailey describes Company, an improvisation ensemble which features a rotating ensemble of improvisers who do not perform regularly together, or as a set group, what he refers to as a "semi-ad-hoc" groupings.[246] Bailey's idea was to focus on the early stages of group development, what he considered to be the most stimulating period, before the musical identity hardens and the music "becomes susceptible to self-analysis, description and, of course, reproduction."[247] However, as Nicholas Cook argues, even when notated chamber scores are performed there is a notion that the performers are listening and negotiating in a social, and much the same process as improvisers.[248] Although I knew the corpora very well, and perhaps I was hoping that the MA would respond with similar slap-tongue material, I did not know how the MA would respond, and if it did respond with similar material, how long it would last. The important detail for me is that in that moment I *trusted* in myself and the MA. I trusted in the material that I was playing, and that there would be a meaningful response from the agent. I also trusted in the global musical result that would occur, maintaining my own responsibility in the musicking, but also allowing space for the MA to take responsibility too. This was both a sense of wanting to try something new, to commit to a specific idea, and to see how the MA would respond, and what proposition or response the MA would make in return. There was also the hope that by committing to a specific idea, we might create a longer structure, rather than abrupt and shorter sectionalized passages which often result. I owe some of this improvement (or

---

[246] Bailey, *Improvisation*, 133.

[247] Bailey, 133.

[248] Nicholas Cook, *Beyond the Score: Music as Performance* (Oxford University Press, 2014), 224–87, https://doi.org/10.1093/acprof:oso/9780199357406.001.0001.

development) to the fact that I knew the corpora and had experience with the MA's past behaviour, although this was a significant experience that had not happened before. In addition, I specifically decided on the material to play in the moment, knowing that I was hoping for a certain kind of musical result (i.e., a polyrhythmic slap-tongued phrase). The results were surprising to me, and to Cusson and Pocius who were involved in the recording session and performance. I question the extent to which my own listening, and my own intention and commitment to playing with the agents was partially responsible for this serendipitous moment. Anne Bogart and Tina Landau instruct artists in theatre to "trust in *letting something occur* onstage, rather than *making it occur*" (italics in original).[249] Was I *trusting* more, or letting go and interacting with the MA, rather than trying to drive or direct the agent? For that matter, what would it even mean to trust in an MA?

6.11 Conclusion

As I have suggested throughout this final chapter, the eTube project and working with MAs has been an incredibly fulfilling experience as a person, artist, and researcher. I have been very privileged to have worked with such dedicated and inspiring individuals who have challenged not only my ways of musicking, but of researching, thinking, creating, writing, and being. Reflecting on this project, I have also considered what other possibilities may have come to fruition during my doctoral research. Sticking with my original topic of live electronic music, I may have ended up with a similar type of document as presented here. I would have commissioned new works, premiered these pieces, collaborated with technicians or programmers on the electronics for performances, and undertaken analyses of these works. This most likely would have resulted in a dissertation including philosophical perspectives of musical interaction as well. However, it was the significant technological hurdles of working with MAs that drove me to reach out to Cusson prior to my second-year recital. It was also the technological hurdles that inspired me to continue working with additional collaborators with specific skills like

[249] Anne Bogart and Tina Landau, *The Viewpoints Book: A Practical Guide to Viewpoints and Composition* (New York: Theatre Communications Group, 2005), 19.

spatialization, which would advance the project in certain directions. In the process of researching artistic and technological solutions is where I am reminded of the wonderful people that I have had the opportunity to collaborate with over the past five years. And it is in this collaborative co-creative environment based around improvisation and a specific philosophical stance towards the MAs which has resulted in a distinctly different kind of co-collaborative project than I might have otherwise undertaken in working on repertoire for live electronics. In this sense I view the MAs and associated computer technologies, as frustrating as they may be at times, as the catalyst that has necessitated this kind of tightly knit interdisciplinary team approach. Although certain people may indeed possess all the programming, performance, and technical skills on their own, it is because of this generosity and commitment from my colleagues that I can personally see the ways these MAs have indeed extended each of our individual agencies and a collective team agency.

I employ the metaphors of the intentional stance and fictional personae to help engage with MAs and to interact in ways that are compelling and meaningful. Reflecting on the propositions above, I ask myself whether I might have insights about what I consider the desires and beliefs of the agents in real-time when improvising? Although I have no clear answer to this question, I would start by suggesting that the attribution of belief or desire is present during the working process of the eTube team as we develop the agent, which in a certain sense, involves "writing our own shared fiction." In sculpting the MA's behaviour through programming and interactive functions in the software, we are indeed affecting the MA's behaviour, and then interacting with these various iterations, and evaluating the result based on our own artistic subjectivities. The adjustments we make to the corpora, software settings, or performance gestures, I suggest, are guided by the conception in our minds of what we would want the agent to believe and desire, which affects their behaviour in performance. In this sense, we share in the creation of our own fictional character who is influenced and shaped by the beliefs and desires of each team member.

In 1964, Marshall McLuhan reminded us in his book *Understanding Media* that "a conscious computer would still be one that was an extension of our consciousness."[250] McLuhan's

[250] McLuhan, *Understanding Media*, 384.

statement seems to broadly assume a general extension of consciousness. One must then ask whose consciousness exactly is being extended, as technologies always involve power dynamics and issues of representation. Rather, I consider that the MAs are a kind of extension specifically of the eTube team's consciousness and agency. One may also point out that by engaging with these technologies like Max/MSP we might also be extending the consciousness of the global programming community, institutions such as IRCAM, and companies such as Cycling '74, including certain shared biases across these communities, some of which were briefly mentioned in Section 6.3 above. In collectively improvised music each performer takes a shared responsibility for the music produced, as described by the core tenet of Tracey Nicholls' "ethos of improvisation."[251] David Borgo posits that group improvisation may result in "complex and emergent properties that are…greater than the sum of its parts."[252] Borgo goes on to suggest that swarm intelligence, as demonstrated by certain insects such as honeybees, has many similarities with improvisation, and may lead to emergent qualities such as nonlinear effects.[253] Although swarm intelligence research is still in its infancy it points towards notions of distributed perception and memory, which are both relevant to how I have framed the MAs as co-collaborative partners throughout this document, and the idea of offloading certain responsibilities on the MAs, as mentioned in Section 6.10.[254] And so, if we are able to engage with MAs in improvisation through imagining make-believe personae, in the same way that one engages with a human improviser, and this might result in complex emergent qualities shown by swarm intelligence, including distributed perception and memory. Then might my interactions and improvisations with MAs also result in the complex and emergent properties, suggesting that the MAs are indeed extending some kind of intentionality and/or agency on behalf of the eTube team members through automatization in performance.

---

[251] Tracey Nicholls, "Speaking Justice, Performing Reconciliation: Twin Challenges for a Postcolonial Ethics," *Critical Studies in Improvisation/Études Critiques En Improvisation* 6, no. 1 (2010): 1.

[252] Borgo, *Sync or Swarm, Revised*, 173.

[253] Borgo, 182–91.

[254] Borgo, 189.

At this point I would like to return to Patrick Burkart's warning stated in Section 4.2.3 above of a "technological utopianism" movement where one "mistakenly places technology, and not human agency, at the source of human history-making."[255] Burkart's concerns regarding misuse of technology, technological biases, and the controlling structures that govern media technologies are indeed concerning. It now seems that AI technologies and autonomous systems are now shaping humans to a greater extent, and so one must consider how certain machine agencies may also play a role in the future of history making. As stated throughout this document, the eTube project shows how human agency is indeed at the centre of creation with MAs, but that these MAs may act to extend some kind of communal human agency. This delicate balance between understanding the human role in the creation process is complicated by the agencies that the MAs may have, or at the very least, the agencies that we attribute to them. It seems more suitable that we consider a both-and situation rather than an either-or in this discussion on human versus machine agency, while keeping in mind the specific contexts these technological agencies play a role in, and which situations they strictly do not. After all, these MAs and associated technologies are often modelled after human cognition and behaviour which inevitably maintain certain biases inherited from their creators. However, whether these MA technologies may actually be *equivalent* to or *replace* humans in performance seems to be an open question. But before even asking that question, one also needs to specify what we mean by "equivalent," and also what it would mean to "replace" a human. In the meantime, what does exploring improvisation with MAs teach us? These technologies seem to be in an intermediary stage, which will undoubtedly evolve significantly as AI research and tools evolve in the coming years. If the MAs do indeed extend a communal agency, one way to conceptualize of the relationship between humans and MAs would be like holding up a mirror to our improvisation practice and reflecting on the desires and beliefs that the team and other collaborators hold. I am reminded that have worked with primarily three MAs presented here, although we have investigated many others to various degrees. As already discussed above, the models of machine "listening" and machine "learning," are exactly that, they are one specific model of these processes, and not necessarily representative of the varied experiences humans have of listening and learning. There are an incredible number of artists creating and performing with MAs. In the same way that there

[255] Burkart, *Music and Cyberliberties*, 121.

should not be one human perspective to emulate, there is no ideal or optimal improvisation, and there is no optimal MA or improvisation with a MA. However, as I hope this project is shown, by whittling down and questioning the artistic and technological decisions throughout this process, the MAs help to teach us more about ourselves as artists and people, in addition to learning about our collaborators in this regard as well.

Appendix A

2021 North American Saxophone Alliance (NASA) Region 10 Conference Program.

NASA NORTH AMERICAN SAXOPHONE ALLIANCE

Region 10 Virtual Conference

January-May, 2021

McGill University Day Professor Marie-Chantal Leclair, host

NOTE: 2 pm EST start: Zoom link and information is on page 2 of this document

The afternoon will be in two parts.

Part 1 (approximately 60-70 minutes)

McGill Saxophone Studio performances

Elliot Carter, Canonic Suite (Fanfare) for four alto saxophones (2')
Robert Fieldhouse, Ella Sandin, Madelyn Carter, Marie-Chantal Leclair

-*Compulsion-Spirale* (2019, rev. 2020) for Alto Saxophone and Live Electronics by Jonas Regnier (11,30)
Tommy Davis, alto saxophone

-*Robert Fieldhouse : Partita (Allemande), J-S Bach*

-*DYCI2* Project (Dynamics of Creative Improvised Interaction) [2017-…] (10')
Marie-Chantal Leclair, saxophone, Tommy Davis, saxophone, Vincent Cusson, electronics

-A brief introduction to *Chaleurs* by Walter Boudreau, for saxophone quartet
Excerpts provided by the Quasar saxophone quartet

(program continues on page 2)

Appendix B

McGill Doctoral Recital 2 Program.

McGill
Schulich School of Music
École de musique Schulich

Salle Tanna Schulich Hall
527, rue Sherbrooke Ouest, Montréal, QC
www.mcgill.ca/music

Le mercredi 26 mai 2021
à 17h

Wednesday, May 26, 2021
5:00 p.m.

Récital de doctorat

Doctoral Recital

Tommy Davis

saxophones / saxophone

classe de / class of Marie-Chantal Leclair

Vincent Cusson, électronique / electronics

Tango[2]	Henning Berg (n. en / b. 1954)
DROP for solo improviser and electronics	Linda Bouchard (n. en / b. 1957)
Construction III	Sergio Kafejian (n. en / b. 1967)
DYCI2	Jérôme Nika (n. en / b. 1988)

Nick Zoulek, saxophone
Kevin Gironnay, électronique / electronics
Marilène Provencher-Leduc, flûte et pédales d'effets numériques / flute and effects pedals

entracte

_derivations for baritone saxophone and interactive performance system	Benjamin Carey (n. en / b. 1984)
Ma	Kevin Gironnay (n. en/ b. 1989)

Kevin Gironnay, électronique / electronics
Marilène Provencher-Leduc, flûte et pédales d'effets numériques / flute and effects pedals

Ce concert sera webdiffusé sur la chaîne YouTube de Schulich
This concert will be webcast on Schulich's YouTube channel
bit.ly/TSHWebcast

Ce concert fait partie des épreuves imposées aux étudiants pour l'obtention de leur diplôme respectif.
This concert is presented in partial fulfilment of the requirements for the degree or diploma programme of the student listed.

Appendix C

McGill Doctoral Lecture-Recital Program.

McGill
Schulich School of Music
École de musique Schulich

Salle Tanna Schulich Hall
527, rue Sherbrooke Ouest, Montréal, QC
www.mcgill.ca/music

Le samedi 25 mars 2023
à 17h

Saturday, March 25, 2023
5:00 p.m.

Conférence-récital de doctorat

Doctoral Lecture-Recital

Tommy Davis

saxophone

classe de / class of
Marie-Chantal Leclair

Tommy Davis, eTube
Kasey Pocius, électronique / electronics, spatialisation

eTu{d,b}e Improvisation Framework	
Creative Dynamics of Improvised Interaction (DYCI2)	Jérôme Nika (n. en / b. 1988)
Construction III	Sergio Kafejian (n. en / b. 1967)
eTu{d,b}e Improvisation Framework	
Spire Muse	Notto J.W. Thelle (n. en / b. 1974)
Enfants, apprenez-nous à parler	Quentin Lauvray (n. en / b. 1997)

Quentin Lauvray, électronique / electronics

Ce concert fait partie des épreuves imposées aux étudiants pour l'obtention de leur diplôme respectif.
This concert is presented in partial fulfilment of the requirements for the degree or diploma programme of the student listed.

Appendix D

Program note for *Enfants, apprenez-nous à parler* (2022) by Quentin Lauvray.

Enfants, apprenez-nous à parler (Children, teach us how to speak) is a piece for eTube and live electronics. The piece is based on an article by Brandt & al. (2012) which shows that the language acquisition in infants is first a musical process. The infant, progressively, playfully, and by a non-linear learning process, selects the sounds they can produce from the chaos of their screams, babbling, comings, and exclamations. Through imitative games and independent explorations, interspersed with the urgent need to express oneself and occasional contemplative silence, the baby learns how to shape phonemes, words, phrases. Based on this idea, the piece is organized in different « learning steps » in which recognizable sounds are extracted from complex and semi-improvised textures. Those sounds gather and are arranged to create more complex configurations and musical phrases. The electronics, which include an improvising agent, act like an external stimulus (a parent) creating pressure on the evolution of the musical discourse, insisting on certain sounds, but which is not insensitive to the musician's propositions.

Appendix E

Program note for *Improvisation Frameworks* (2022) by Greg Bruce and Tommy Davis.

Greg Bruce and Tommy Davis, in collaboration with Kasey Pocius and Vincent Cusson, have developed several improvisation frameworks for feedback saxophone, eTube, and improvising digital agents. For their debut concert at *live@CIRMMT* in December 2022, the duo assembled the following approaches: 1) acoustic limits duo; 2) feedback agent leading; 3) water tube agent trio; and 4) multi-agent crescendo. For these two artist-researchers, improvisation is a key method in systematizing and expanding their respective electroacoustic systems, and therefore is a natural setting for combining the two instruments for the first time. In contrast with the duo's solo works, these frameworks explore and develop inter-system interactions and investigate the flexibility of each instrument in improvised dialogues.

Appendix F

Program note for *3tube* (2023) by Kasey Pocius.

Building upon Kasey and Tommy's 2022 eTube performance produced by Codes d'accès, 3tube features the Weather Vane Collective with three eTube performers and improvising software. Each tube is outfitted with a different mouthpiece, enabling a distinct timbral palette for each performer. In this collaborative composition, Pocius has expanded the musical vocabulary of the agents to complement each musician's personality and existing musical practice. This includes incorporating Pocius's field recording and modular synthesizer practice.

The piece is divided into 8 sections, each focusing on a method of interaction between the ensemble, the agents and the fixed media. This piece also serves as a way to continue developing techniques such as timbral transfer between sound sources, spatialization, creating comprovisional systems to complement individual performers. This work also follows my previous *Piano Dreamscape* pieces, which looked at ways to record improvisations on acoustic instruments in such a way as to capture their acoustic radiation patterns while utilizing close mics to capture details, then using spectral and granular processes on these recordings to emphasize the materiality of acoustic instruments.

This piece utilizes CCCP, DYCI2 and Construction III.

Fixed media draws inspiration from
Harrison, Going/Places
Truax, Island
Carey, Contingent States
Lee, Teum (the Silvery Slit)
Moth Cock, Castles Off Jersey

Bibliography

Albrecht, Robert, Jussi Jaatinen, Tapio Lokki, and others. "Electronic Hearing Protection for Musicians." In *Proceedings of the 14th Sound and Music Computing Conference*, 306–13, 2017. https://urn.fi/URN:NBN:fi:aalto-201708036338.

Alissa Cheung, Violin. "Alissa Cheung, Violin." Accessed March 28, 2024. http://www.alissacheung.com/.

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

Assayag, G., and S. Dubnov. "Using Factor Oracles for Machine Improvisation." *Soft Computing* 8, no. 9 (2004): 604–10. https://doi.org/10.1007/s00500-004-0385-4.

Assayag, Gérard, Georges Bloch, and Marc Chemillier. "Omax-Ofon." In *Sound and Music Computing*, 2006. https://hal.science/hal-01161346.

Backus, John. "Effect of Wall Material on the Steady-State Tone Quality of Woodwind Instruments." *The Journal of the Acoustical Society of America* 36, no. 10 (1964): 1881–87. https://doi.org/10.1121/1.1919286.

Bailey, Derek. *Improvisation: Its Nature and Practice in Music*. New York: Da Capo Press, 1993.

Bauer, Valentin, Dimitri Soudoplatoff, Leonard Menon, and Amandine Pras. "Binaural Headphone Monitoring to Enhance Musicians' Immersion in Performance." In *Advances in Fundamental and Applied Research on Spatial Audio*, edited by Brian F. G. Katz and Piotr Majdak, 193–219. Rijeka: IntechOpen, 2022. https://doi.org/10.5772/intechopen.104845.

Bhagwati, Sandeep. "Virtuosities of the Native Alien." In *Contemporary Musical Virtuosities*, edited by Louise Devenish and Cat Hope, 68–78. Routledge Research in Music. London: Routledge, 2024. https://doi.org/10.4324/9781003307969.

Bogart, Anne, and Tina Landau. *The Viewpoints Book: A Practical Guide to Viewpoints and Composition*. New York: Theatre Communications Group, 2005.

Bonnasse-Gahot, Laurent. "An Update on the SOMax Project." *Ircam-STMS, Tech. Rep*, 2014.

Borgo, David. "Openness from Closure: The Puzzle of Interagency in Improvised Music and a Neocybernetic Solution." In *Negotiated Moments: Improvisation, Sound, and Subjectivity*, edited by Gillian H. Siddall and Ellen Waterman, 113–30. Improvisation, Community, and Social Practice. Durham: Duke University Press, 2016. https://doi.org/10.1515/9780822374497.

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

Born, Georgina. *Rationalizing Culture: IRCAM, Boulez, and the Institutionalization of the Musical Avant-Garde*. Berkeley: University of California Press, 1995.

Bowers, John, and Phil Archer. "Not Hyper, Not Meta, Not Cyber but Infra-Instruments." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 5–10. Vancouver, Canada, 2005. https://www.nime.org/proceedings/2005/nime2005_005.pdf.

Bown, Oliver. "Player Responses to a Live Algorithm: Conceptualising Computational Creativity without Recourse to Human Comparisons?" In *The Sixth International*

Conference on Computational Creativity, 126–33. Park City, UT, 2015.

Bown, Oliver, Arne Eigenfeldt, Aengus Martin, Benjamin Carey, and Philippe Pasquier. "The Musical Metacreation Weekend: Challenges Arising from the Live Presentation of Musically Metacreative Systems." In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 9:27–34, 2013.

Bown, Oliver, Alice Eldridge, and Jon McCormack. "Understanding Interaction in Contemporary Digital Music: From Instruments to Behavioural Objects." *Organised Sound* 14, no. 2 (August 2009): 188–96. https://doi.org/10.1017/S1355771809000296.

Braasch, Jonas. *Hyper-Specializing in Saxophone Using Acoustical Insight and Deep Listening Skills*. Vol. Vol. 6. Current Research in Systematic Musicology. Cham, Switzerland: Springer, 2019. https://doi.org/10.1007/978-3-030-15046-4.

Brown, Andrew R., Matthew Horrigan, Arne Eigenfeldt, Toby Gifford, Daniel Field, and Jon McCormack. "Interacting with Musebots." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 19–24. Blacksburg, VA, 2018.

Bruce, Greg. "Surmounting the Skepticism: Developing a Research-Creation Methodology." *Acta Academiae Artium Vilnensis*, no. 109 (August 2023): 101–23. https://doi.org/10.37522/aaav.109.2023.162.

Bruce, Gregory Andrew. "Feedback Saxophone: Expanding the Microphonic Process in Post-Digital Research-Creation." Doctoral dissertation, University of Toronto, 2023. https://hdl.handle.net/1807/130068.

Burkart, Patrick. *Music and Cyberliberties*. Music Culture. Middletown: Wesleyan University Press, 2012. muse.jhu.edu/book/435.

Cannam, Chris, Christian Landone, and Mark Sandler. "Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files." In *Proceedings of the ACM Multimedia 2010 International Conference*, 1467–68. Firenze, Italy, 2010.

Carey, Benjamin. "_derivations: Improvisation for Tenor Saxophone and Interactive Performance System." In *Proceedings of the 9th ACM Conference on Creativity & Cognition*, 411–12. Sydney, Australia: ACM, 2013. https://doi.org/10.1145/2466627.2481226.

Carey, Benjamin. "Designing for Cumulative Interactivity: The _derivations System." In *Proceedings of the International Conference on New Interfaces for Musical Expression*. University of Michigan, Ann Arbor, MI, 2012. https://doi.org/10.5281/zenodo.1178227.

Carey, Benjamin Leigh. "_derivations and the Performer-Developer: Co-Evolving Digital Artefacts and Human-Machine Performance Practices." PhD thesis, University of Technology, 2016. https://opus.lib.uts.edu.au/handle/10453/43452.

Carried Away: Wild Space Dance & Duo d'Entre-Deux @ Roulette, NYC [Excerpt 1]. YouTube video, 04:46, 2015. https://www.youtube.com/watch?v=INmeuZ3ahl0.

Cavdir, Doga, and Sofia Dahl. "Performers' Use of Space and Body in Movement Interaction with A Movement-Based Digital Musical Instrument." In *Proceedings of the 8th International Conference on Movement and Computing*. New York, NY: Association for Computing Machinery, 2022. https://doi.org/10.1145/3537972.3537976.

Chion, Michel. *Sound: An Acoulogical Treatise*. Translated by James A. Steintrager. Durham: Duke University Press, 2015. https://doi.org/10.1515/9780822374824.

Christin, Angèle. "The Ethnographer and the Algorithm: Beyond the Black Box." *Theory and Society* 49, no. 5–6 (2020): 897–918.

CIRMMT. "live@CIRMMT: Improvising New Winds." Accessed February 28, 2024.

https://www.cirmmt.org/en/events/live-cirmmt/improvising-new-winds.
CIRMMT. “Welcome to CIRMMT.” Accessed March 3, 2024. https://www.cirmmt.org/en.
Clynes, Manfred E., and Nathan S. Kline. “Cyborgs and Space.” *Astronautics* 5, no. 9 (1960): 26–27, 74–76.
Codes d’accès. “Cod’a 2022: Échanges.” Accessed February 29, 2024. https://codesdacces.org/evenement/coda-2022-echanges/.
Codes d’accès. “Cod:A-19.” Accessed March 3, 2024. https://codesdacces.org/evenement/coda-19/.
Codes d’accès. “melting links.” Accessed February 29, 2024. https://codesdacces.org/evenement/melting-links/.
Conduit. *Oxford English Dictionary*. online: Oxford University Press. Accessed January 15, 2023. https://www.oed.com/view/Entry/38645?rskey=SwQMV6&result=1.
Cook, Nicholas. *Beyond the Score: Music as Performance*. Oxford University Press, 2014. https://doi.org/10.1093/acprof:oso/9780199357406.001.0001.
Csikszentmihalyi, Mihaly. *Flow: The Psychology of Optimal Experience*. First edition. New York: Harper & Row, 1990.
Currie, Gregory. *The Nature of Fiction*. Cambridge: Cambridge University Press, 1990. https://doi.org/10.1017/CBO9780511897498.
Cusson, Vincent. “Home.” Vincent Cusson. Accessed February 24, 2024. https://vincentcusson.github.io/.
Cusson, Vincent. “Vincent Cusson/eTube,” April 30, 2023. https://github.com/VincentCusson/eTube.
Cusson, Vincent, and Tommy Davis. “Etu{d,b}e: A Preliminary Conduit.” In *Proceedings of the International Conference on New Interfaces for Musical Expression*. The University of Auckland, New Zealand, 2022. https://doi.org/10.21428/92fbeb44.c05957ee.
“Cycling ’74.” Accessed April 25, 2024. https://cycling74.com/.
Davis, Tommy, Kasey Pocius, Vincent Cusson, Maxwell Gentili-Morin, and Philippe Pasquier. “Embodied eTube Gestures and Agency.” In *Proceedings of the 9th International Conference on Movement and Computing*. New York, NY: Association for Computing Machinery, 2024. https://doi.org/10. 1145/3658852.3659084.
Davis, Tommy, Kasey Pocius, Vincent Cusson, Marcelo M. Wanderley, and Philippe Pasquier. “eTu{d,b}e: Case Studies in Playing with Musical Agents.” In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 268–76. Mexico City, Mexico, 2023. http://nime.org/proceedings/2023/nime2023_39.pdf.
Dennett, Daniel C. *The Intentional Stance*. Cambridge, MA: MIT Press, 1987. https://proxy.library.mcgill.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=48751&scope=site.
Devine, Kyle. *Decomposed: The Political Ecology of Music*. Cambridge, MA: MIT Press, 2019. https://doi.org/10.7551/mitpress/10692.003.0001.
Di Donato, Balandino, Christopher Dewey, and Tychonas Michailidis. “Human-Sound Interaction: Towards a Human-Centred Sonic Interaction Design Approach.” In *Proceedings of the 7th International Conference on Movement and Computing*, 1–4, 2020. https://doi.org/10.1145/3401956.3404233.
“Dicy2 | Ircam Forum.” Accessed March 28, 2024. https://forum.ircam.fr/projects/detail/dicy2/#project-versions-anchor.
“Dissertation Media - YouTube.” Accessed April 30, 2024.

https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.
Dix, Alan. "Human-Computer Interaction." In *Encyclopedia of Database Systems*, edited by Ling Liu and M. Tamer Özsu, 1327–31. Boston, MA: Springer US, 2009. https://doi.org/10.1007/978-0-387-39940-9_192.
Dourish, Paul. *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge, MA: MIT press, 2001.
DPA. "4099 Instrument Condenser Microphone." Accessed March 23, 2024. https://www.dpamicrophones.com/instrument/4099-instrument-microphone.
Duo d'Entre-Deux. "Home." Accessed February 24, 2024. http://www.duodentredeux.com.
DYCI2 by Jérôme Nika - Two Soprano Saxophone and Computer Agent Improvisation. YouTube video, 7:18, 2021. https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.
"DYCI2/Dicy2." Accessed March 28, 2024. https://github.com/DYCI2/Dicy2.
Eigenfeldt, Arne, Oliver Bown, Andrew R. Brown, and Toby Gifford. "Distributed Musical Decision-Making in an Ensemble of Musebots: Dramatic Changes and Endings." In *Proceedings of the International Conference on Computational Creativity*, 88–95, 2017.
Eigenfeldt, Arne, Oliver Bown, Philippe Pasquier, and Aengus Martin. "Towards a Taxonomy of Musical Metacreation: Reflections on the First Musical Metacreation Weekend." *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* 9, no. 1 (2013).
Emmerson, Simon. *Living Electronic Music*. Burlington, VT: Ashgate, 2007.
Ensemble AKA + Louis Beaudoin - de La Sablonnière [No Hay Banda]. YouTube video, 25:15, 2017. https://www.youtube.com/watch?v=9rotwgh_ZTE.
eTube – Spire Muse Sessions | Take 2. YouTube video, 3:11, 2022. https://www.youtube.com/watch?v=49LuS84ZOxw.
eTube – Spire Muse Sessions | Take 5. YouTube video, 3:57, 2022. https://www.youtube.com/watch?v=Rq5HJ07etOI.
eTube – Spire Muse Sessions | Take 6. YouTube video, 5:41, 2022. https://www.youtube.com/watch?v=nbHG0cFNafs.
eTube Techniques Demonstration. YouTube video, 0:55, 2023. https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.
eTu{d,b}e by Kasey Pocius and Tommy Davis – Cod'a 2022: Échanges. YouTube video, 19:40, 2024. https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.
eTu{d,b}e by Kasey Pocius and Tommy Davis – IIICON 2023. YouTube video, 9:56, 2023. https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.
eTu{d,b}e by Vincent Cusson and Tommy Davis. YouTube video, 8:18, 2023. https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.
eTu{d,b}e by Vincent Cusson and Tommy Davis - NIME 2022. YouTube video, 14:12, 2022. https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.
Fairfield, Joshua A. T. "The Magic Circle." *Vand. J. Ent. & Tech. L.* 11 (2009): 823–40.
Federman, Jeremy, and Todd Ricketts. "Preferred and Minimum Acceptable Listening Levels for Musicians While Using Floor and In-Ear Monitors." *Journal of Speech, Language, and Hearing Research* 51, no. 1 (2008): 147–59.
Fletcher, Neville. "Materials for Musical Instruments." *Acoustics Australia* 27, no. 1 (1999): 5–9.
Fletcher, Neville H., and Thomas D. Rossing. *The Physics of Musical Instruments*. 2nd ed. New

York: Springer, 1998.
FLUX:: Immersive. “SPAT Revolution.” Accessed March 23, 2024. https://www.flux.audio/project/spat-revolution/.
Goehr, Lydia. *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*. Revised edition. Oxford: Oxford University Press, 2007. https://ebookcentral.proquest.com/lib/mcgill/detail.action?docID=3053152.
Goldstein, Malcolm. *Sounding the Full Circle: Concerning Improvisation and Other Related Matters*. Sheffield, VT: M. Goldstein, 1988. http://www.frogpeak.org/unbound/index.html.
Goodlad, Lauren M. E. “Editor’s Introduction: Humanities in the Loop.” *Critical AI* 1, no. 1–2 (October 2023). https://doi.org/10.1215/2834703X-10734016.
Greg Bruce Music. “Home.” Accessed February 25, 2024. https://www.gregbruce.ca.
Haraway, Donna. “A Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s.” *Australian Feminist Studies* 2, no. 4 (March 1, 1987): 1–42. https://doi.org/10.1080/08164649.1987.9961538.
Harker, Alexander, and Pierre Alexandre Tremblay. “The HISSTools Impulse Response Toolbox: Convolution for the Masses.” In *Proceedings of the International Computer Music Conference*, 148–55. The International Computer Music Association, 2012. http://eprints.hud.ac.uk/id/eprint/14897/.
Harvey, Paul. *Saxophone*. London: Kahn & Averill, 1995.
Hasegawa, Robert. “Creating with Constraints.” In *The Oxford Handbook of the Creative Process in Music*, edited by Nicolas Donin. Oxford University Press, online edition, 2020. https://doi.org/10.1093/oxfordhb/9780190636197.013.17.
Hayles, N. Katherine. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: The University of Chicago Press, 1999. http://hdl.handle.net/2027/heb.05711.
Improvisation Frameworks by Greg Bruce and Tommy Davis [Premiere]. YouTube video, 10:10, 2024. https://www.youtube.com/playlist?list=PLxU35HjSL4p1sCRF1JeqYMstfpO1yqvVg.
Instruments, Interfaces, Infrastructures: An Interdisciplinary Conference on Musical Media. “Conference Program.” Accessed February 24, 2024. https://sites.harvard.edu/instruments-interfaces-infrastructures/.
Iyer, Vijay. “Improvisation: Terms and Conditions.” In *Arcana IV: Musicians on Music*, edited by John Zorn, 171–75. New York: Hips Road, 2009.
Kafejian, Sergio. Personal communication with the author, January 15, 2023.
Kaiser, Jeff. “Improvising Technology, Constructing Virtuosity.” *Cuadernos de Música, Artes Visuales y Artes Escénicas* 13, no. 2 (2018): 87–96.
Kasey Pocius. “Kasey Pocius.” Accessed March 23, 2024. https://www.kaseypocius.ca.
KMC Music. “Home.” Accessed March 23, 2024. https://kmcmusic.com/.
Koetsier, Teun. “On the Prehistory of Programmable Machines: Musical Automata, Looms, Calculators.” *Mechanism and Machine Theory* 36, no. 5 (May 1, 2001): 589–603. https://doi.org/10.1016/S0094-114X(01)00005-2.
Kramer, Lawrence. “Philosophizing Musically: Reconsidering Music and Ideas.” *Journal of the Royal Musical Association* 139, no. 2 (2014): 387–404.
Lamarque, Peter, and Stein Haugom Olsen. *Truth, Fiction, and Literature: A Philosophical Perspective*. Oxford University Press, 1996.

https://doi.org/10.1093/acprof:oso/9780198236818.001.0001.
Lee, Rosemary. "The Limits of Algorithmic Perception: Technological Umwelt." In *Politics of the Machines-Art and After*, 1–6. Aalborg University, Copenhagen, Denmark: BCS Learning & Development, 2018. https://doi.org/10.14236/ewic/EVAC18.44.
Legault, Maryse. "Music." Maryse Legault. Accessed February 25, 2024. https://maryselegault.com.
Leman, Marc. *Embodied Music Cognition and Mediation Technology*. Cambridge, MA: MIT Press, 2008. https://doi.org/10.7551/mitpress/7476.001.0001.
Lessig, Lawrence. *Remix: Making Art and Commerce Thrive in the Hybrid Economy*. 1st ed. London: Bloomsbury Academic, 2008. https://doi.org/10.5040/9781849662505.
Levinson, Jerrold. "Musical Expressiveness as Hearability-as-Expression." In *Contemporary Debates in Aesthetics and the Philosophy of Art*, edited by Matthew Kieran, 192–206. Malden, MA: Blackwell, 2006.
Lévy, Benjamin, Georges Bloch, and Gérard Assayag. "OMaxist Dialectics." In *Proceedings of the International Conference on New Interfaces for Musical Expression*. University of Michigan, Ann Arbor, 2012. https://hal.science/hal-00706662/.
Lewis, Eric. *Intents and Purposes: Philosophy and the Aesthetics of Improvisation*. Ann Arbor: University of Michigan Press, 2019. https://proxy.library.mcgill.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=2046031&scope=site.
Lewis, George E. "Co-Creation: Early Steps and Future Prospects." In *Artisticiel / Cyber-Improvisations*, edited by Bernard Lubat, Gérard Assayag, and Marc Chemillier. Dialogiques d'Uzeste. Phonofaune, 2021. https://hal.science/hal-03542917.
Lewis, George E. "Improvised Music after 1950: Afrological and Eurological Perspectives." In *The Other Side of Nowhere: Jazz, Improvisation, and Communities in Dialogue*, edited by Daniel Fischlin and Ajay Heble, 1st ed., 131–62. Music/Culture. Middletown, CT: Wesleyan University Press, 2004. http://catdir.loc.gov/catdir/toc/ecip048/2003019192.html.
Lewis, George E. "Too Many Notes: Computers, Complexity and Culture in 'Voyager.'" *Leonardo Music Journal* 10 (2000): 33–39. https://www.jstor.org/stable/1513376.
Lewis, George E. "Why Do We Want Our Computers to Improvise?" In *The Oxford Handbook of Algorithmic Music*, edited by Roger T. Dean and Alex McLean, Vol. 1. Oxford University Press, 2018. https://doi.org/10.1093/oxfordhb/9780190226992.013.29.
Luminous - Tubes Excerpt - Wild Space Dance Company & Duo d'Entre-Deux. YouTube video, 1:22, 2015. https://www.youtube.com/watch?v=G0XABGUDFfA.
Malloch, Joseph, and Marcelo Wanderley. "Embodied Cognition and Digital Musical Instruments: Design and Performance." In *The Routledge Companion to Embodied Music Interaction*, edited by Micheline Lesaffre, Pieter-Jan Maes, and Marc Leman, 1st ed., 438–47. New York: Routledge, 2017. https://doi-org.proxy3.library.mcgill.ca/10.4324/9781315621364.
Mamedes, Clayton Rosa, A. Rodrigues, Marcelo M. Wanderley, Jônatas Manzolli, Denise HL Garcia, and Paulo Ferreira-Lopes. "Composing for DMIs - Entoa, Music for Intonaspacio." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 509–12. Goldsmiths, University of London, UK, 2014.
Masu, Raul, Fabio Morreale, and Alexander Refsum Jensenius. "The O in NIME: Reflecting on the Importance of Reusing and Repurposing Old Musical Instruments." In *Proceedings of*

the International Conference on New Interfaces for Musical Expression. Mexico City, Mexico, 2023.
"Maxwell Gentili-Morin – IDMIL." Accessed March 26, 2024. https://www.idmil.org/people/maxwell-gentili-morin/.
McCormack, Jon, Toby Gifford, Patrick Hutchings, Maria Llano, Matthew Yee-King, and Mark d'Inverno. "In a Silent Way: Communication Between AI and Improvising Musicians Beyond Sound," 1–11, 2019. https://doi.org/10.1145/3290605.3300268.
McLuhan, Marshall. *Understanding Media: The Extensions of Man*. 1964. Reprint, London: Routledge, 2001.
"MIDImix." Accessed March 23, 2024. https://www.akaipro.com/midimix.
Millman, Matthew. "Matt's Tech Pages." *Common JST Connector Types* (blog), January 7, 2021. https://www.mattmillman.com/info/crimpconnectors/common-jst-connector-types/.
Miranda, Eduardo Reck, and Marcelo M. Wanderley. *New Digital Musical Instruments: Control and Interaction beyond the Keyboard*. Middleton, Wisconsin: A-R Editions, Inc., 2006. http://site.ebrary.com/id/10857388.
Mirandau, Eduardo Reck. *A-Life for Music: Music and Computer Models of Living Systems*. Edited by Eduardo Reck Miranda. Computer Music and Digital Audio Series; Volume 24. Middleton, Wisconsin: A-R Editions, 2011. https://ebookcentral.proquest.com/lib/mcgill/detail.action?docID=3115105.
Monson, Ingrid. *Saying Something: Jazz Improvisation and Interaction*. Chicago: University of Chicago Press, 1996. http://www.SLQ.eblib.com.au/patron/FullRecord.aspx?p=432268.
Morreale, Fabio, and Andrew McPherson. "Design for Longevity: Ongoing Use of Instruments from NIME 2010-14." In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Aalborg University Copenhagen, Denmark, 2017.
"Music Representations Team." Accessed February 24, 2024. http://repmus.ircam.fr/impro.
Musical Metacreation. "About," November 12, 2015. https://musicalmetacreation.org/about/.
Nicholls, Tracey. "Speaking Justice, Performing Reconciliation: Twin Challenges for a Postcolonial Ethics." *Critical Studies in Improvisation/Études Critiques En Improvisation* 6, no. 1 (2010): 1–15.
Nick Zoulek. "Saxophonist | Composer." Accessed February 24, 2024. http://nickzoulek.com.
Nika, Jérôme, Marc Chemillier, and Gérard Assayag. "ImproteK: Introducing Scenarios into Human-Computer Music Improvisation." *ACM Computers in Entertainment* 14, no. 2 (2017): 1–27. https://doi.org/10.1145/3022635.
Nika, Jérôme, Ken Déguernel, Axel Chemla–Romeu-Santos, Emmanuel Vincent, and Gérard Assayag. "DYCI2 Agents: Merging the 'Free', 'Reactive', and 'Scenario-Based' Music Generation Paradigms." In *International Computer Music Conference*. Shanghai, China, 2017. https://hal.archives-ouvertes.fr/hal-01583089.
Nika, Jérôme, Augustin Muller, Joakim Borg, Gérard Assayag, and Matthew Ostrowski. "Dicy2 for Max." Ircam UMR STMS 9912, December 2022. https://hal.science/hal-03892611.
Østern, Tone Pernille, Sofia Jusslin, Kristian Nødtvedt Knudsen, Pauliina Maapalo, and Ingrid Bjørkøy. "A Performative Paradigm for Post-Qualitative Inquiry." *Qualitative Research* 23, no. 2 (April 2023): 272–89.
Oxford English Dictionary. "Collaborator, n., Sense 1." Oxford University Press, July 2023. Oxford English Dictionary. https://doi.org/10.1093/OED/5793806289.
Oxford English Dictionary. "Partner, n.1, Sense 2.a." Oxford University Press, December 2023. Oxford English Dictionary. https://doi.org/10.1093/OED/1044963391.

Oxford English Dictionary. “Semantics.” Oxford University Press, July 2023. Oxford English Dictionary. https://doi.org/10.1093/OED/2159347396.

Oxford English Dictionary. “Syntax.” Oxford University Press, September 2023. Oxford English Dictionary. https://doi.org/10.1093/OED/1187492139.

Paine, Garth. “Interactivity, Where to from Here?” *Organised Sound* 7, no. 3 (2002): 295–304.

Pasquier, Philippe, Arne Eigenfeldt, Oliver Bown, and Shlomo Dubnov. “An Introduction to Musical Metacreation.” *Computers in Entertainment* 14, no. 2 (December 2016): 2:1-2:14. https://dl.acm.org/doi/10.1145/2930672.

Perez-Osorio and Wykowska. “Adopting the Intentional Stance toward Natural and Artificial Agents.” *Philosophical Psychology* 33, no. 3 (2020): 369–95. https://doi.org/10.1080/09515089.2019.1688778.

“Physics Makerspace.” Accessed March 22, 2024. https://makerspace.physics.mcgill.ca/printers.html.

Pocius, Kasey. “Expanding Spatialization Tools for Various DMIs.” Master’s Thesis, McGill University, 2023.

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

Pras, Amandine, Michael F. Schober, and Neta Spiro. “What about Their Performance Do Free Jazz Improvisers Agree upon? A Case Study.” *Frontiers in Psychology* 8 (2017): 1–19. https://doi.org/10.3389/fpsyg.2017.00966.

Pressing, Jeff. “Psychological Constraints on Improvisational Expertise and Communication.” In *In the Course of Performance: Studies in the World of Musical Improvisation*, edited by Bruno Nettl and Melinda Russell, 47–67. Chicago Studies in Ethnomusicology. Chicago, IL: University of Chicago Press, 1998.

Puckette, Miller. “Max at Seventeen.” *Computer Music Journal* 26, no. 4 (2002): 31–43. http://www.jstor.org.proxy3.library.mcgill.ca/stable/3681767.

“RE20 - Broadcast Announcer’s Microphone with Variable-D by Electro-Voice.” Accessed April 15, 2024. https://products.electrovoice.com/na/en/re20/.

Reverberant House: Presented by Duo d’Entre-Deux - YouTube. YouTube video, 01:00. Accessed February 24, 2024. https://www.youtube.com/watch?v=uKuuD6CkHbk.

Ridenour, Tom. “The Grenadilla Myth,” no date. https://www.rclarinetproducts.com/the-grenadilla-myth.

Rowe, Robert. *Interactive Music Systems: Machine Listening and Composing*. Cambridge, MA: MIT Press, 1993.

Rowe, Robert. “Machine Listening and Composing with Cypher.” *Computer Music Journal* 16, no. 1 (1992): 43–63. https://doi.org/10.2307/3680494.

Rowe, Robert. *Machine Musicianship*. Cambridge, MA: MIT Press, 2001.

“Sam Newsome - Home.” Accessed March 23, 2024. http://www.somenewmusic.com/.

Sam Newsome - Solo Soprano Saxophone Live on WFMU - July 13, 2018. YouTube video, 52:00, 2018. https://www.youtube.com/watch?v=hKyetZ42nn4.

Sanden, Paul. *Liveness in Modern Music: Musicians, Technology, and the Perception of Performance*. New York, NY: Routledge, 2017.

Schacher, Jan C. “Moving Music: Exploring Movement-to-Sound Relationships.” In *Proceedings of the 3rd International Symposium on Movement and Computing*. New

York, NY: Association for Computing Machinery, 2016. https://doi.org/10.1145/2948910.2948940.
Schier, Flint. "Tragedy and the Community of Sentiment." In *Philosophy and Fiction*, edited by Peter Lamarque, 73–92. Aberdeen: Aberdeen University Press, 1983.
Schmidtmann, Gunnar, Andrew J. Logan, Claus-Christian Carbon, Joshua T. Loong, and Ian Gold. "In the Blink of an Eye: Reading Mental States From Briefly Presented Eye Regions." *I-Perception* 11, no. 5 (2020). https://doi.org/10.1177/2041669520961116.
Schulich School of Music. "Doctoral Recital: Thomas Davis, Saxophone." Accessed March 3, 2024. https://www.mcgill.ca/music/channels/event/doctoral-recital-thomas-davis-saxophone-329499.
Schwarz, Diemo. "Concatenative Sound Synthesis: The Early Years." *Journal of New Music Research* 35, no. 1 (March 1, 2006): 3–22. https://doi.org/10.1080/09298210600696857.
Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–24.
Siddall, Gillian, and Ellen Waterman. "Introduction: Improvising at the Nexus of Discursive and Material Bodies." In *Negotiated Moments: Improvisation, Sound, and Subjectivity*, edited by Gillian Siddall and Ellen Waterman. Improvisation, Community, and Social Practice. Durham: Duke University Press, 2016. https://doi.org/10.1515/9780822374497.
"Sirnotto/Cccp: Co-Creative Communication Platform." Accessed March 26, 2024. https://github.com/sirnotto/cccp.
Small, Christopher. *Musicking: The Meanings of Performing and Listening*. Music/Culture. Hanover: University Press of New England Hanover, 1998.
Sone, Yuji. "Dead Stars and 'Live' Singers: Posthumous 'Holographic' Performances in the US and Japan." In *Sound and Robotics*, edited by Richard Savery, 317–36. Chapman and Hall/CRC, 2022. https://doi.org/10.1201/9781003320470.
"Spat | Ircam Forum." Accessed March 3, 2024. https://forum.ircam.fr/projects/detail/spat/.
Stenros, Jaakko. "In Defence of a Magic Circle: The Social, Mental and Cultural Boundaries of Play." *Transactions of the Digital Games Research Association* 1, no. 2 (2014).
Sterne, Jonathan. "Is Machine Listening Listening?" *Communication +1* 9, no. 1 (October 2022). https://doi.org/10.7275/zeqh-eg38.
Sterne, Jonathan, and Elena Razlogova. "Machine Learning in Context, or Learning from LANDR: Artificial Intelligence and the Platformization of Music Mastering." *Social Media + Society* 5, no. 2 (2019). https://doi.org/10.1177/2056305119847525.
Stévance, Sophie, and Serge Lacasse. *Research-Creation in Music and the Arts: Towards a Collaborative Interdiscipline*. Sempre Studies in the Psychology of Music. New York: Routledge, 2018.
Tatar, Kıvanç, and Philippe Pasquier. "MASOM: A Musical Agent Architecture Based on Self Organizing Maps, Affective Computing, and Variable Markov Models." In *Proceedings of the 5th International Workshop on Musical Metacreation*, 2017.
Tatar, Kıvanç, and Philippe Pasquier. "Musical Agents: A Typology and State of the Art towards Musical Metacreation." *Journal of New Music Research* 48, no. 1 (January 1, 2019): 56–105. https://doi.org/10.1080/09298215.2018.1511736.
Thelle, Notto J. W. "Mixed-Initiative Music Making: Collective Agency in Interactive Music Systems." Doctoral dissertation, Norges musikkhøgskole, 2022.
Thelle, Notto J. W., and Philippe Pasquier. "Spire Muse: A Virtual Musical Partner for Creative Brainstorming." In *Proceedings of the International Conference on New Interfaces for*

Musical Expression. NYU Shanghai, Shanghai, China, 2021. https://doi.org/10.21428/92fbeb44.84c0b364.

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

Tremblay, Pierre Alexandre, Nicolas Boucher, and Sylvain Pohu. "Real-Time Processing on the Road: A Guided Tour of [Iks]'s Abstr/Cncr Setup." In *International Computer Music Conference*, 2007. http://eprints.hud.ac.uk/id/eprint/999/.

Uexküll, Jakob von. "A Stroll through the Worlds of Animals and Men: A Picture Book of Invisible Worlds." *Semiotica* 89, no. 4 (1992): 319–91. https://doi.org/10.1515/semi.1992.89.4.319.

Wanderley, Marcelo M., Bradley W Vines, Neil Middleton, Cory McKay, and Wesley Hatch. "The Musical Significance of Clarinetists' Ancillary Gestures: An Exploration of the Field." *Journal of New Music Research* 34, no. 1 (2005): 97–113.

Wang, Johnty, Axel Mulder, and Marcelo M. Wanderley. "Practical Considerations for MIDI over Bluetooth Low Energy as a Wireless Interface." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 25–30. Porto Alegre, Brazil, 2019.

Weiss, Marcus, and Giorgio Netti. *The Techniques of Saxophone Playing*. Kassel: Bärenreiter, 2010.

Wiggins, Geraint A. "A Preliminary Framework for Description, Analysis and Comparison of Creative Systems." *Knowledge-Based Systems* 19, no. 7 (2006): 449–58. https://doi.org/10.1016/j.knosys.2006.04.009.

Winkler, Todd. *Composing Interactive Music: Techniques and Ideas Using Max*. Cambridge, MA: MIT Press, 1998. https://proxy.library.mcgill.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1430&scope=site.

Wooldridge, Michael. "Intelligent Agents." In *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, edited by Gerhard Weiss, 27–78. Cambridge, MA: MIT Press, 1999. http://cognet.mit.edu/library/books/view?isbn=0262731312.

Young, Michael. "NN Music: Improvising with a `Living' Computer." In *Computer Music Modeling and Retrieval. Sense of Sounds*, edited by Richard Kronland-Martinet, Sølvi Ystad, and Kristoffer Jensen, 337–50. Berlin: Springer, 2008. https://doi.org/10.1007/978-3-540-85035-9_23.

Young, Michael, and Tim Blackwell. "Live Algorithms for Music: Can Computers Be Improvisers?" In *The Oxford Handbook of Critical Improvisation Studies, Volume 2*, edited by Benjamin Piekut and George E. Lewis, 507–28. Oxford: Oxford University Press, 2016. https://doi.org/10.1093/oxfordhb/9780199892921.013.002.

Zunshine, Lisa. "Why We Read Fiction." *Skeptical Inquirer* 30, no. 6 (2006): 29–33.