eTu{d,b}e: case studies in playing with musical agents

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

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

Vincent Cusson
IDMIL,CIRMMT
McGill University
Montreal, Canada
vincent.cusson2@mail.mcgill.ca

Marcelo M. Wanderley
IDMIL,CIRMMT
McGill University
Montreal, Canada
marcelo.wanderley@mcgill.ca

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

ABSTRACT

The *eTu{d,b}e* framework adapts existing improvising musical agents (MA) for performance with an augmented instrument called the eTube. This instrument has been developed with deliberate musical and technological limitations including a simple two-button controller and restricted pitch capacity. We will present case studies which outline our research-creation framework for mapping the eTube controller, developing corpora for the MAs, and testing interactive and machine listening settings which will also be demonstrated by performance examples. A general summary of the MAs will be followed by specific descriptions of the features we have utilised in our work, and finally a comparison of the MAs based on these features. Few papers discuss the process for learning to work with and adapt existing MAs and we will finish by describing challenges experienced as other users with these technologies.

Author Keywords

Musical agents, augmented instrument, improvised performance

CCS Concepts

•**Applied computing** → **Sound and music computing;** Performing arts; •**Human-centered computing** → Interaction design;

1. INTRODUCTION

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 [13]. We present a simple plastic tube, a proto-wind instrument that has been augmented and is our conduit for studying improvising musical agents. A conduit for sharing our love of improvisation and computer music. A conduit for structuring our improvisation practice. A conduit for testing bi-directional interaction with improvising agents. A conduit for imposing performance limitations on an improviser, providing clear boundaries within which to explore creatively [18]. This research-creation project is built upon a simple and frugally designed cylindrical instrument which is our conduit to share our process of learning, adapting, and performing with existing musical agents designed by other developers.

Plastic tubes have been used to extend saxophones while maintaining traditional performance techniques (Sam Newsome 2018; Intended Assembly 2018). Saxophonists Tommy Davis and Nick Zoulek have performed on saxophone mouthpieces and necks attached to various plastic tubes as the Duo d'Entre-Deux in improvised and dance performances since 2014 (Carried Away 2015; Luminous 2015; Reverberant House 2019).[1] These experiences laid the foundation for Davis' tube practice integrating movement, spatialised performance gestures, and contemporary techniques.

The eTube is an augmented instrument made of a 2.54 cm diameter, 219 cm long cylindrical polyvinyl chloride (PVC) tube augmented with a simple controller and fitted with a baritone saxophone mouthpiece. This purposefully limited proto-wind instrument produces an intriguing sound world, although it has no keys or tone holes to change pitch. These restrictions are inspired by "infra-instruments," which are purposefully incomplete, deconstructed, or broken instruments with electronic augmentation [4]. The eTube's sound is less timbrally rich and its performance techniques are limited compared to traditional acoustic instruments, which allows more latitude for electronic processing [4]. Acoustically, the instrument's harmonics resemble the overtone series with certain partials absent and non-typical tuning for other partials (Figure 1). This limited sound world lends itself well to "non-idiomatic" [2, 25] improvisation contexts, where musical properties other than Western European melodic, harmonic, and rhythmic development are foregrounded [25].[2]

The eTube controller (Figure 2) is designed by instrument builder and programmer Vincent Cusson in collaboration with Tommy Davis and uses available low-cost electronics including 3D printed parts to anchor the controller to the tube.[3] The simple controller design is inspired by the saxophone's keys and right-hand thumb rest. The instrument

[1] https://www.duodentredeux.com
[2] https://youtu.be/AH3lvAdAwiQ (eTube performance example)
[3] https://github.com/VincentCusson/eTube

NIME'23, 31 May–2 June, 2023, Mexico City, Mexico.

Figure 1: eTube overtones accessible by overblowing, transcribed by Quentin Lauvray

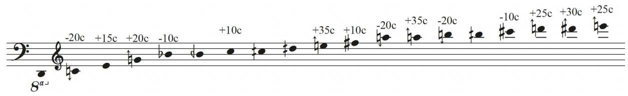
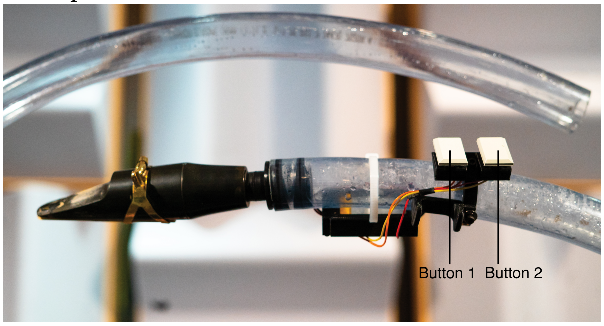


Figure 2: eTube with controller and baritone saxophone mouthpiece attached



is supported by the performer's left hand, reproducing a saxophonist's hand position, which aims to maintain instrumental technique and to minimize fatigue while holding the eTube directly in one's hand (see Figure 3). A custom 3D printed mouthpiece adapter designed by Cusson ensures the mouthpiece is securely connected to the eTube. The parametric 3D plan is adaptable for other purposes and allows the parts to be adjusted for various-sized mouthpieces and tube diameters, promoting versatility.

Figure 3: Davis performing the eTube

The two-button controller is designed to facilitate interaction with improvising musical agents (MA) [34]. MA research is part of Musical Metacreation, a subfield of Computational Creativity, which "focuses on endowing machines with the ability to achieve creative musical tasks, such as...improvisation" [29]. MAs consist of real-time synthesis algorithms and an interface with sensor technology mapped to these algorithms allows a performer to communicate with

Controller Action	Intended Interaction
Single click button 1	Query agent 1
Double click button 2	Turn on/off agent 2 listening
Long click both buttons	Start/stop loop recorder
Double click both buttons	Start/stop loop playback

Table 1: ***DYCI2*** **and** ***Construction III*** **mapping table**

the MAs [22]. The buttons produce a click when depressed and provide auditory feedback for the performer. The controller modeled on the saxophone's mechanisms has been an intuitive solution for Davis to communicate with agents in performance.

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

Table 2: ***Spire Muse*** **mapping table**

The controller was originally conceived to map global adjustments in existing MAs; however, mappings are currently limited to direct commands, such as querying the agent or turning on or off listening settings (see Tables 1 and 2). The two-button controller limits the number of mappings for any performance. We embrace this limitation and change the mappings for different performances depending on the corpora, collaborators, performance context [5], and the MA(s) used.

Although patch settings may be manipulated by a technician in real-time via the user interface (UI), the eTube controller enables the performer to adjust these settings during performance. This allows the performer another level of communication with the MAs besides auditory feedback. The controller design aimed to provide Davis with additional interactive capabilities with the agents, while freeing Cusson from managing the MAs onstage (see Section 4 below).

We are grateful to the collaborators mentioned in this paper, who have introduced us to novel approaches to musicking with MAs [32]. The eTube shapes, guides, and informs our research-creation approach to performing with the MAs and adapting *eTud,be* framework. Our 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. We assess the musical and artistic outcomes of the improvisations from the performers' perspective and continue to explore different artistic possibilities that emerge through interactions with MAs during performances.

1.1 The eTu{d,b}e Framework

First developed by Cusson and Davis, the *eTu{d,b}e* framework adapts existing MAs in a flexible performance architecture utilising one or multiple MAs as described in Section 3. *eTu{d,b}e* simultaneously refers to the name of the eTube and to a series of improvised etudes based on human-computer musical interactions. The French word *étudier* (to study) suggests that the performer and MAs *étudient* (study) each other in performance. The performer and/or programmer learn how agents react differently with certain corpora or listening settings, and the agent studies the performer's sound using audio descriptors and machine listening.

eTu{d,b}e performances investigate interactive elements such as listening presets, corpus analysis, eTube controller mappings, and the number of agents. Our evaluations are undertaken by the programmer/performer team and are based on how our mappings and corpora facilitate interaction with MAs, and the resultant performance. These MAs are implemented in Max/MSP[4], allowing adaptation for specific performance situations or controller specifications. The programmers and improvisers are continually learning about the musical agent through testing and performance [31]. However, real-time learning functions for the MAs have yet to be implemented in our framework [**?**, 28].

2. RELATED WORK

Many developers have created and documented their own MAs for improvised performance such as Robert Rowe [30], George E. Lewis [21], Michael Young [37], and Benjamin Carey [11]; however, few papers describe new augmented instruments being adapted for performance with existing MAs and the associated collaborative team, working process, and artistic output.

George E. Lewis' *Voyager* [21] is known for its skilled and musically appropriate output which is based on algorithms defined by Lewis. *Voyager* is autonomous, it uses MIDI messages to generate musical material and does not need a performer's input to function. Although *Voyager* is using MIDI, the performer is only reacting with the agent's audio output. The rule-based algorithms defined by Lewis result in a specific improvised output and the agent does not learn new styles [34]. Rather than using representational MIDI data, our MAs undergo an audio corpora training phase which includes audio segmentation, audio descriptor analysis, and a clustering process to label the audio segments. The programmer and/or improviser curate the corpora to vary in style and musical material. The MA's output utilise machine listening based on audio descriptors and requires a performer's input to generate musical material.

Jon McCormack et al. [23] have used visual stimuli to communicate the MA's confidence to the performer using emoticons displayed on a screen in addition to auditory feedback. This multimodal communication between human and agent was shown to enhance performer flow states [15] during performance. The visual feedback may also contribute to improvised performances being more musically interesting, with clearer balance between human and agent as perceived by audiences listening to audio recordings [23]. We use lighting automated to the amplitude of the MA's output as a visual feedback mechanism to aid audience members to distinguish between acoustic sounds and agent's output.

The Musical Metacreation Weekend featured concerts of MAs with performers and subsequently Oliver Bown et al. [5] discuss reasons why MA reproducibility has not been widely adopted. Published papers often do not contain sufficient documentation to reproduce MAs. Even after multiple performances it may still be challenging to evaluate MA performance, and in certain cases, the creator may not have complete understanding of their MA's output. These obstacles contribute to difficulties for other users.

3. MUSICAL AGENT DESCRIPTIONS

The *eTu{d,b}e* framework utilises three existing MAs including the *Creative Dynamics of Improvised Interaction* (*DYCI2*) by Jérôme Nika et al. [27], *Spire Muse* by Notto Thelle et al. [35], and *Construction Tools for Interactive Performance* (*CTIP*) by Sergio Kafejian. *CTIP* has been updated by Cusson for *eTu{d,b}e* and this new version is referred to as *Construction III*. These MAs were chosen because of the developers' willingness to collaborate with us. We initially began with *DYCI2* and *Construction III*, the former being a more interactive and the latter a more reactive system, before adding *Spire Muse* [34]. The *eTu{d,b}e* framework is adaptable and we often combine multiple MAs in performance with the eTube. In addition, distinct controller mappings and corpora are implemented for different MAs and performances (see Section 4).

DYCI2 and *Spire Muse* are corpus-based MAs and *Construction III* employs real-time processing modules [34]. The training process for *DYCI2* and *Spire Muse* first segments sound recordings, then uses audio descriptors for analysis, and finally, a clustering process creates audio labels for each segment. *Construction III* employs multiple agents which undertake listening, analysis, and sound generation tasks. Rather than using corpora, *Construction III* routes incoming audio through the effects modules which are processed in real-time [19].

3.1 DYCI2

DYCI2 is built on *Omax* [20], *Somax* [3], and *Improtek* [26] and enables multiple agents, each with separate corpora and listening settings [27]. Agent training takes place offline, and one must train each agent individually by clicking through the menu or using presets. Many audio descriptors are included in the training menu, which can be intimidating and poses challenges for evaluating the agent's behaviour if descriptors are continually changed between performances.

Response settings for the agents are selected on the UI (Figures 4 and 5) 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. Each agent has an adjustable threshold value, which is the amplitude needed to emit a musical response from the agent.

3.2 Spire Muse

Spire Muse is built upon the *Musical Agent Based on Self-Organizing Maps* (MASOM) architecture [33] and employs one large corpus. 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. Once a model is created from the corpus, it is easily dragged and dropped into the interface (Figure 6) 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 [35].

3.3 Construction III

Construction III (Figure 7) is a performance-driven environment that utilises transformative and sequenced techniques [30] organised into various effects processing modules including a multitrack recorder with fixed and randomized

[4]https://cycling74.com/

Figure 4: DYCI2 interface—one agent

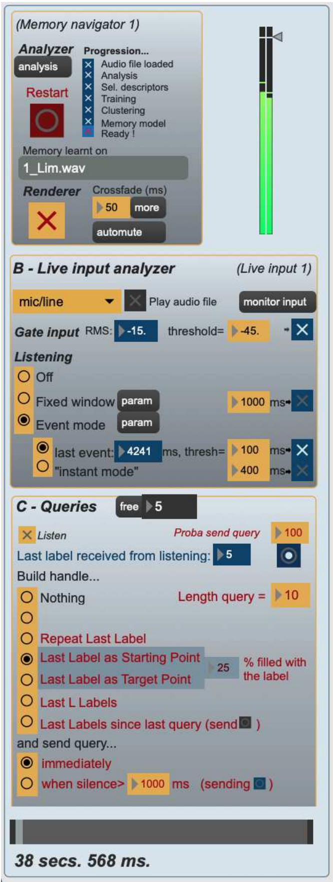
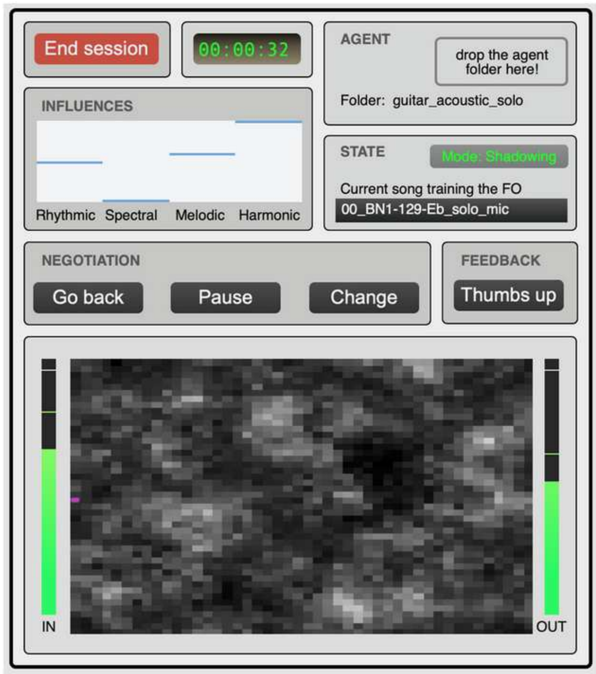


Figure 5: DYCI2 interface—three agents

Figure 6: Spire Muse interface [35]



variable-speed playback, reverberation, four-channel delay, granulator, and custom spectral processing effects [19]. The MA has a built-in spatialisation module with outputs in stereo, quadraphonic, or octophonic [19].

A customisable matrix routes the performer's input to the effects modules based on a pitch tracker with three separate and adjustable ranges. The performer may interact with the agent by performing material in one of the three ranges, thus affecting how the audio is routed to one or more effects modules. These rule-based routing processes result in a MA that is closer to the reactive end of the reactive-autonomous continuum [34].

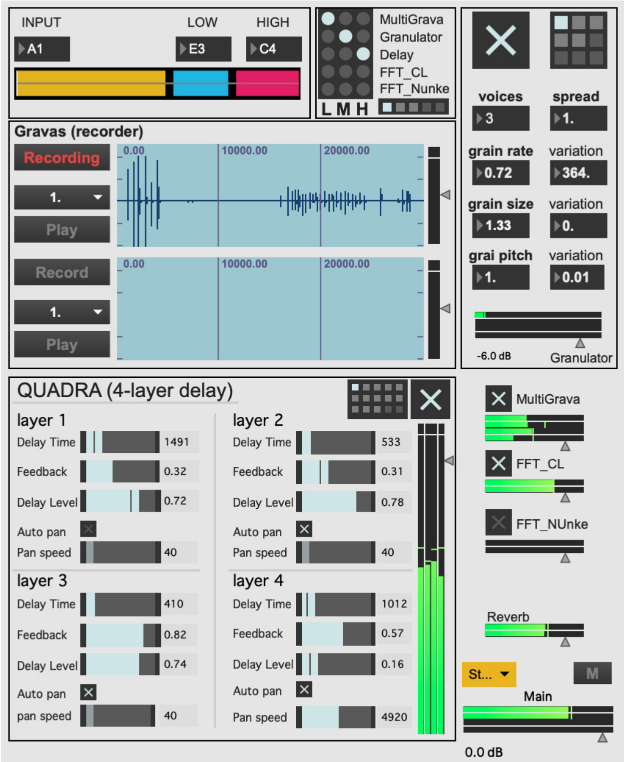
3.4 Research-Creation Framework

The eTube is the basis for developing our research-creation framework involving performing with and adapting existing MAs for performance with the eTube controller. This project is an interdisciplinary undertaking and encompasses fields like music performance, technology, instrument design, and programming. Using this research-creation framework, the eTube controller mappings and corpora are performed, tested, and adapted for each MA. The performance limitations mentioned in Section 1 also influence the controller design and mappings, corpora recordings, and evaluation process.

Once an initial version of the controller was operational, the research-creation process followed a general outline consisting of planning, implementation, testing, and reworking stages. The project directions were determined in part by upcoming performances and based on an ongoing testing and reflection process. Most of the updates include controller mappings, corpora implementation, and adapting various interactive settings in the MAs. Each MA offers possibilities for interaction which are inherent in the settings and UI. New changes were implemented and tested in rehearsals and performances which were often recorded. The team would analyse and reflect upon the musical result and either adopt, abandon, or continue to adapt the new updates. Tools to capture real-time controller and agent data to inform our reflection process are being developed and will be discussed in Section 5.2.

Like the infra-instrument concept mentioned in Section 1, most aspects of this project have been approached with similar limitations in mind. Corpora curation has been limited primarily to tube sounds and saxophone recordings and all corpora presented in the case studies have been recorded by Davis (see Section 4). The team could have used numerous saxophone recordings for the corpora, but wished to keep the eTube as an extension of Davis' ongoing exploration of the tube's sonic identity. We were familiar with the improvised material in the MA's corpora because we had recorded

Figure 7: Construction III interface



it ourselves, and this helped us to evaluate the MA's listening settings. Using corpora that were homogenous sounding with the eTube resulted in a blurred perception of the sounds produced acoustically and electronically. This also problematizes the distinction between an instrument that is augmented with electronics, and performance of an instrument with MAs as two separate realities. Recording and testing our corpora has been a continual re-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 Davis continues to advance performance techniques. Similar co-creative processes between performer and agents have been discussed by Notto Thelle [35] and Benjamin Carey [10].

4. PERFORMANCE CASE STUDIES

> "Live performance and improvisation are amongst the most challenging creative activities undertaken by humans." [23]

4.1 NIME 2022 Video Performance

eTu{d,b}e was premiered at NIME 2022 via an online video recorded at the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT) by Cusson and Davis in May 2022 [1].[5] Three *DYCI2* agents and *Construction III*, which were both adapted by Cusson, were used for this performance. *DYCI2*'s agent one featured an eTube slap-tongue articulation 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 (Figure 2) twice consecutively would turn agent two on or off (Table 1). 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 some were triggered with a lower volume input from Davis and other agents required a higher volume input. These threshold settings helped to create varied responses which influenced the form of the improvisation and allowed Davis to sonically trigger certain agents with a lower volume input, while evoking a musical response from all the agents with a louder input.

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. Simultaneously holding both buttons started the recording, and it stopped when the two buttons were released. A double click of both buttons would launch the recording playback (Table 1). This module has not been routed to *DYCI2* to interact with the agents and remains primarily an expressive tool for the performer to reinforce, interact with, or to suggest formal structure by using playback of past improvised material.

During the recording process Cusson and Davis noticed that they were starting to hear the agent's output in new ways which could be described as algorithmic listening [7], where they were able to understand certain elements or settings based on the agent's responses (or lack thereof) in performance via the sonic output. At times this helped to speed up troubleshooting as the two had more insight into which setting, agent, or corpora to adjust. Davis also felt more confident with the agents in performance since he could hear specific settings reflected in the agent's behaviour.

4.2 Metacreation Lab Residency

In spring 2022 Davis and Cusson undertook a CIRMMT Inter-centre Research Exchange at Simon Fraser University's (SFU) Metacreation Lab supervised by Prof. Philippe Pasquier.[6] The two investigated and tested *Spire Muse* and adapted the eTube controller to interact with specific settings. For the *Spire Muse* corpus, the team used acoustic tube recordings from the NIME 2022 recording session mentioned in Section 4.1. A series of performances with *Spire Muse* recorded by Davis concluded the SFU residency.[7]

The eTube controller was programmed so a double-click of button one toggled the "Change" function, a double click of button two activated the "Go back" function, and pressing both buttons simultaneously initiated the "Pause" operation (see Table 2). The "Change" function advances the influences to new values, while "Go back" reverts to the preceding settings. Davis used these functions to suggest form in the improvisation by introducing new material with the "Change" function, then revisiting previous ideas after toggling "Go back."

4.3 Bruce-Davis CIRMMT Recording

In June 2022 Cusson and Davis recorded with Greg Bruce performing on his feedback saxophone, marking the first eTube collaboration with another instrument.[8] Bruce's feedback saxophone system includes a guitar amplifier, analogue guitar pedals, and a tenor saxophone with a DPA lavalier microphone fixed inside the bell. The microphone cre-

[5] https://www.youtube.com/watch?v=TN8i9gqPuKc

[6] https://metacreation.net

[7] https://youtu.be/49LuS84ZOxw

[8] https://www.gregbruce.ca

ates feedback with the amplifier and depending on the saxophone fingering used, different feedback pitches are heard. Bruce maintains technical facility on the saxophone and performs acoustically with these feedback tones.

Spire Muse was used with the same eTube corpus employed for the SFU residency mentioned in Section 4.2. During the session, Bruce commented that the eTube and MA were in competition against his feedback saxophone because the MA and eTube sounds were similar in timbre and gesture. In response, Davis did not use the controller, and a saxophone-only corpus was implemented to differentiate the MA from the eTube. This resulted in the sense of a trio improvisation including eTube, feedback saxophone, and a MA [8].

A video recording of this performance implemented an intensity effect automated to the MA's loudness which was superimposed over a Genelec speaker's power light, creating a flicker effect when the MA performed.[9] Issues of disembodied electronics in performance with acoustic instruments has been discussed by John Croft [14] and Simon Emmerson [17] among others. We chose amplitude for this lighting effect because we believed it to be the most meaningful parameter for the audience and we hoped the visual indicator would help listeners distinguish between agent and human interjections.

4.4 Codes d'Accès 2022 Performance

For the Codes d'accès performance Kasey Pocius worked with Davis and Cusson to rework the *eTu{d,b}e* NIME 2022 patch to add collaborative quadraphonic spatialisation, more hands-on mixing of the MA's outputs, and further control of the processing from *Construction III*.[10] Additionally, Pocius suggested a "dog-bone" structure where all the agents and processing begin at full volume, which then quickly decay and are slowly reintroduced throughout the piece until all processing is removed at the end, allowing the acoustic eTube to shine.

Three stereo outputs from the *DYCI2* agents plus 18 channels from *Construction III* were spatialised throughout the quadraphonic system using Spat5 for Max/MSP [12]. This patch was programmed such that Pocius could manually interpolate between four spatial presets using their Akai MidiMix [11], while an envelope follower on the eTube mic was used to interpolate between adjacent presets. The MIDI controller faders were used to control the MA's levels and the processing. Davis continued to have control over the MAs and processing via the controller (see Section 4.1).

A recording of the eTube filled partially with water was used for *DYCI2*'s agent one. Agent two consisted of a short low drone. Davis recorded a new corpus with articulated, isolated harmonics for agent three which helped to add rhythmic variation and responses.

4.5 Enfants, Apprenez-Nous à Parler

Enfants, apprenez-nous à parler (*Children, teach us to speak*) (2022) by Quentin Lauvray is the first commissioned piece for the eTube and explores motherese and baby talk as a metaphor for the expressive but limited proto-instrument qualities of the eTube.[12] The work features composed and improvised sections using spatialised fixed audio and five *DYCI2* agents. Lauvray utilized Cusson's version of *DYCI2* described in Section 4.1 with spatialisation work by Pocius in Spat5 for Max/MSP. The eTube controller was used to trigger fixed audio cues in the Max patch.

The composition develops as a metaphor for how infants are thought to learn a language, first via rhythm, timbre, and melody, followed by syntax and meaning [6]. The work begins with simple rhythmic and melodic fragments which develop into longer phrase structures throughout. For the corpora, Davis improvised on the eTube to recordings of infant and mother communicating. Lauvray manually segmented and categorized these recordings by hand based on the performance technique and sonic gesture, resulting in corpora consisting of short fragments. Lauvray used these specific corpora to limit the agent's behaviour, helping to maintain the work's trajectory and form [31]. Three dynamic microphones were used to trigger *DYCI2* agents and for processing the eTube. Horizontal eTube movements across the microphone array were indicated in the score and were related to the electronic's spatialised trajectories.

4.6 Other eTube Users

Greg Bruce is the first performer other than Davis to perform with the eTube and *eTu{d,b}e* framework. Bruce supplied his own mouthpiece setup, differing from Davis'. An initial acoustic performance allowed Bruce to become accustomed to the eTube's performance capabilities without MAs. Bruce explored moving the eTube while playing, investigating the instrument's flexible nature and directional sound. *DYCI2* was used with the same corpora employed for the performance in Section 4.4 and the controller was programmed as outlined in Section 4.1. Bruce began performing with each agent one by one, while Pocius and Davis adjusted the MAs' settings such as threshold values, agent response length, and delaying the agent response following a dialogue with Bruce. For the final performances with all three MAs, one additional controller mapping allowed Bruce to turn on or off the water tube corpus.

Gestures played an important role in how Bruce interacted with the agents. He played with the proximity of the eTube to the microphone, interacting with the volume threshold at which an agent's response was triggered. Bruce used the controller to guide the improvisation, triggering the agents to respond to his sonic gestures and turning agents on and off to influence density [9].

4.7 Musical Agent Comparison

The following comparisons are from the performers' perspective and attempt to highlight certain inherent features of the MAs which we explored through performance.

DYCI2 has the most complex UI with many machine listening and interactive settings which provides the user with considerable control over the parameters for analysis and synthesis. However, similar agent behaviour may be achieved with different setting combinations, which may cause evaluation and reproducibility difficulties. MA's behaviours were discovered each time parameters, corpus, or playing technique changed, and so the authors often reverted to parameters that had most recently worked, often relying on pitch chroma and amplitude. *Spire Muse* has a simple interface, and when in interactive mode, the influences change automatically allowing a user to perform with musical results after training the corpus. *Construction III* incorporates effects modules, and although there are many adjustable settings, we had musically-satisfying performances more quickly than with the corpora-based MAs.

In *DYCI2* the same corpus may be used to train multiple agents but with different analysis descriptors, result-

[9] https://youtu.be/JjChhV5tPSw
[10] https://youtu.be/n97wToOFiJo
[11] https://www.akaipro.com/midimix
[12] https://youtu.be/5Otow16e7dk

ing in two or more agents with distinct analyses of the same corpus. Implementing multiple agents trained this way during a performance would not be possible with a mono-agent setup. Each *DYCI2* agent has an adjustable threshold value. These values can be useful to structure the density of the agents, setting certain thresholds to be triggered easily (low volume input) and others to necessitate a high input (high volume input).

When first working with *Spire Muse*, Davis noted that the MA's responses were unique and varied. At times the *DYCI2* agents would repeat the same segment multiple times in performances, which seemed to occur less with *Spire Muse*. The eTube's limited sound production and the homogenous-sounding corpora may be a constraining factor here. This difference might also be attributable to *Spire Muse*'s larger corpus compared with the smaller corpora for each *DYCI2* agent. This is not to suggest that the corpus size is the sole determining factor for a musically successful performance, however. *Spire Muse*'s interactive modes and influences change automatically which may also aid to vary the agent's response. The repeated segments from *DYCI2*, although a noticeable difference when compared to *Spire Muse*, was not necessarily a limiting factor in terms of musicality. Much of what we learned with *DYCI2*, including developing corpora, was applied to our work with *Spire Muse* and may have decreased the time it took to successfully perform with the latter.

5. LIMITATIONS AND FUTURE WORK

> "Programming a patch for live electronics is always a matter of compromise between what is technologically possible and what is likely to crash the computer." [24]

5.1 Other User Problems

DYCI2 and *Spire Muse* are open source software and *Construction III* was received directly from Sergio Kafejian. These MAs have not been rigorously tested on multiple operating systems and have one person, or small teams, making updates. These projects are often not well documented, and issues routinely need to be discussed with the developer to be solved or circumvented using alternative programming. A lack of documentation is an ongoing issue for developers without dedicated support teams and is understandable due to the time commitment necessary to develop these resources. When support documentation is available, there remains a steep learning curve for other users due to the complexity of the MA's processes [5]. Regardless of these challenges, documentation is nonetheless helpful for other users learning MAs. In addition, email correspondence and online meetings with the developers have been integral to advancing our research.

The eTube presents challenges to the listener because it is a new instrument with little cultural relevance or established performance practice [16]. All sound manipulation is concealed in the mouth and there are no external keys to suggest physical correlation with the sonic output. In addition, the corpora used for MAs in performances are primarily acoustic eTube improvisations by Davis. Audiences may have difficulty understanding what sounds are acoustic and which are electronic since they are derived from the same instrument. This may affect their comprehension of computer agency and interaction throughout the performance; however, blurring the boundaries between instrument and machine can be a useful artistic tool [24].

Based on the GitHub documentation, the eTube hardware should be reproduceable by other artists with access to a 3D printer and the electronic devices, although the authors are not aware of anyone who has reproduced the eTube.[13] While someone could potentially build an eTube, the challenges related to learning and working with the $eTu\{d,b\}e$ framework, including curating corpora, manipulating settings, and evaluating the agent's output [5] may present more of a barrier for artists to successfully adapt or perform with this system [22].

The eTube is capable of limited sounds when using traditional playing methods but contemporary techniques [36] expand these possibilities, although they take years to refine. The eTube is light, flexible, and directional, and spatialising the sound with performance gestures has proven to be an intuitive performance technique. The highly gestural and specialised performance techniques, including technical challenges relating to the MAs, may be a limiting factor for adoption of this instrument by other artists. These challenges may also inspire other users to develop innovative solutions, opening up new directions for improvisation with digital musical instruments (DMI).

It was anticipated that the eTube's limited sonic nature would also aid in evaluating the machine listening and interaction between human and agents due to the relatively homogeneous corpora and acoustic input being implemented (see eTube performance example 2). Further testing is required to confirm or disprove this claim.

5.2 Visualisation Tools

Visualisation tools are being developed to collect controller data which will aid our analysis of the communication between performer and MA(s). A timestamp and label for each controller input are synchronized with the recording in a digital audio workstation (DAW). This will enable us to consider how the controller mappings affect the resulting performance. These data have been collected for multiple performances, but no analysis has been undertaken.

6. CONCLUSIONS

We have presented the $eTu\{d,b\}e$ framework and the eTube augmented instrument, which are our inspiration and method for learning and adapting existing MAs in improvised performance. This research includes the programmer/instrument builder-performer/improviser relationship and the details learned through our research-creation process over the past three years.

We have introduced three MAs used in $eTu\{d,b\}e$ and how we have adapted these agents for performance with the eTube and controller. This process has included developing and testing interactive controller mappings, corpora curation, and testing audio descriptors and machine listening settings.

Five case studies outline the specific ways that our research-creation process has evolved on an artistic level while discussing how we have implemented the mappings, corpora, and listening settings mentioned above. Preliminary ideas were shared regarding other users performing the eTube and visualisation tools for data collection related to the controller. Although many challenges exist when working with MAs created by other developers, we believe that our research provides one context for working with, learning from, and performing alongside musical agents.

[13] https://github.com/VincentCusson/eTube

7. ACKNOWLEDGMENTS

The eTube and $eTu\{d,b\}e$ have been developed thanks to Student Research Awards (2021-22, 2022-23) and Intercentre Research Exchange Funding from the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT) at McGill University. This research is also supported by the Social Sciences and Humanities Research Council of Canada (SSHRC), a Natural Sciences and Engineering Research Council of Canada (NSERC) grant supporting the Input Devices and Musical Interaction Laboratory (IDMIL) at McGill, the Metacreation Lab at Simon Fraser University, the Canada Council for the Arts, and *Conseil des Arts de Montréal*. We would like to thank our supervisors Robert Hasegawa, Marie-Chantal Leclair, and Marcelo M. Wanderley for their support and guidance during this research. The ideas for this paper were solidified during a CIRMMT Inter-centre Research Exchange at Simon Fraser University's Metacreation Lab in spring 2022 which was hosted and supervised by Philippe Pasquier. A special thank you to Philippe for his commitment and support of this research during the SFU residency and throughout the writing process. Thank you to Jérome Nika, Sergio Kafejian, and Notto Thelle the programmers and composers who generously shared their knowledge of their improvising systems through email and online meetings. Thank you to Quentin Lauvray for collaborating on the first composition for the eTube. Thank you to Ben Carey and Henning Berg for sharing their knowledge of their MAs during our early performances which informed the eTube development. Thanks to Yves Methot at CIRMMT for the support during the early motion capture & research recordings. Thank you to Richard McKenzie at the Digital Composition Studio (DCS) at McGill for his ongoing support throughout the project.

8. ETHICAL STANDARDS

The authors do not recognize any potential conflicts of interest in this research project. Artistic collaborators and consultants who were engaged for the project were compensated at rates at or above market as determined by the relevant union representing their creative practice.

Having been enormously inspired by various projects made in the NIME community, we think it's important to share this ongoing work with an open-source license. The use of off-the-shelf electronics, hardware, and 3D-printed parts should make this project more accessible and replicable. All source code and documentation to reproduce the instrument are made available via this public GitHub repository. We encourage people to use, share and modify this work to fit into their own creative practices.

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

9. REFERENCES

[1] Cusson, Vincent and Davis, Tommy. *Etu{d,b}e*: A Preliminary Conduit. In *Proceedings of the 2022 conference on New Interfaces for Musical Expression*, 2022. https://nime.pubpub.org/pub/wvhyx5ch.

[2] D. Bailey. *Improvisation: Its nature and practice in music*. Da Capo Press, New York, NY, 1993.

[3] L. Bonnasse-Gahot. An update on the SOMax project. Technical report, IRCAM, 2014.

[4] J. Bowers and P. Archer. Not Hyper, Not Meta, Not Cyber but Infra-Instruments. In *Proceedings of the 2005 conference on New Interfaces for Musical Expression*, 2005.

[5] O. Bown, A. Eigenfeldt, A. Martin, B. Carey, and P. Pasquier. The Musical Metacreation Weekend: Challenges Arising from the Live Presentation of Musically Metacreative Systems. In *In Proceedings of the Artificial Intelligence and Interactive Digital Entertainment Conference*, pages 27–34, 2013.

[6] A. Brandt, M. Gebrian, and L. R. Slevc. Music and Early Language Acquisition. *Frontiers in Psychology*, 3, 2012.

[7] A. R. Brown, M. Horrigan, A. Eigenfeldt, T. Gifford, D. Field, and J. McCormack. Interacting with musebots. In *Proceedings of the 2018 conference on New Interfaces for Musical Expression*, pages 19–24, 2018.

[8] G. Bruce. Personal communication with Vincent Cusson and Tommy Davis, Jun 28, 2022.

[9] G. Bruce. Personal communication with Kasey Pocius and Tommy Davis, Jan 17, 2023.

[10] B. Carey. Designing for cumulative interactivity: The _derivations system. In *Proceedings of the 2012 conference on New Interfaces for Musical Expression*, 2012.

[11] B. Carey. _derivations: Improvisation for Tenor Saxophone and Interactive Performance System. In *Proceedings of the 9th ACM Conference on Creativity & Cognition*, pages 411–412, 2013.

[12] T. Carpentier, M. Noisternig, and O. Warusfel. Twenty years of ircam spat: looking back, looking forward. In *Proceedings of the 2015 International Computer Music Conference*, pages 270–277, 2015.

[13] Conduit. *Oxford English Dictionary*. Oxford University Press, online. Accessed Jan 15, 2023 from https://www.oed.com/view/Entry/38645?rskey=SwQMV6&result=1.

[14] J. Croft. Theses on liveness. *Organised Sound*, 12(1):59–66, 2007.

[15] M. Csikszentmihalyi. *Flow: The psychology of optimal experience*. Harper & Row, New York, 1990.

[16] S. Emmerson. Acoustic/electroacoustic: The relationship with instruments. *Journal of New Music Research*, 27(1-2):146–164, 1998.

[17] S. Emmerson. 'Losing touch?': the human performer and electronics. In *Music, Electronic Media and Culture*, pages 194–216. Ashgate, Burlington, VT, 2000.

[18] R. Hasegawa. Creating with Constraints. In N. Donin, editor, *The Oxford Handbook of the Creative Process in Music*. Oxford University Press, online edition, Jul 2020.

[19] S. Kafejian. Personal communication with Tommy Davis, Jan 15, 2023.

[20] B. Lévy, G. Bloch, and G. Assayag. OMaxist Dialectics. In *Proceedings of the 2012 conference on*

New Interfaces for Musical Expression, 2012.

[21] G. E. Lewis. Too many notes: Computers, complexity and culture in voyager. *Leonardo Music Journal*, 10:33–39, 2000.

[22] J. Malloch and M. M. Wanderley. Embodied cognition and digital musical instruments: design and performance. In *The Routledge companion to embodied music interaction*, pages 438–447. Routledge, 2017.

[23] J. McCormack, T. Gifford, P. Hutchings, M. Llano, M. Yee-King, and M. d'Inverno. In a Silent Way: Communication Between AI and Improvising Musicians Beyond Sound. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–11, 2019.

[24] S. McLaughlin. If a tree falls in an empty forest . . . : Problematization of liveness in mixed-music performance. *Journal of Music, Technology & Education*, 5(1):17–27, May 2012.

[25] J. Morris. *Perpetual Frontier: The Properties of Free Music*. Riti Publishing, Stony Creek, CT, 2012.

[26] J. Nika, M. Chemillier, and G. Assayag. ImproteK: Introducing scenarios into human-computer music improvisation. *ACM Computers in Entertainment*, 14(2):1–27, 2017.

[27] J. Nika, K. Déguernel, A. Chemla–Romeu-Santos, E. Vincent, and G. Assayag. DYCI2 agents: Merging the "free", "reactive", and "scenario-based" music generation paradigms. In *International Computer Music Conference*, 2017.

[28] G. Paine. Interactivity, where to from here? *Organised Sound*, 7(3):295–304, 2002.

[29] P. Pasquier, A. Eigenfeldt, O. Bown, and S. Dubnov. An Introduction to Musical Metacreation. *Computers in Entertainment*, 14(2):1–14, Dec. 2016.

[30] R. Rowe. Machine Listening and Composing with Cypher. *Computer Music Journal*, 16(1):43, 1992.

[31] R. Rowe. *Machine musicianship*. MIT press, 2004.

[32] C. Small. *Musicking: The meanings of performing and listening*. Wesleyan University Press, 1998.

[33] K. Tatar and P. 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 (MUME 2017)*, pages 1–8, 2017.

[34] K. Tatar and P. Pasquier. Musical agents: A typology and state of the art towards Musical Metacreation. *Journal of New Music Research*, 48(1):56–105, Jan. 2019.

[35] N. J. W. Thelle and P. Pasquier. Spire Muse: A Virtual Musical Partner for Creative Brainstorming. In *Proceedings of the 2021 conference on New Interfaces for Musical Expression*, 2021.

[36] M. Weiss and G. Netti. *The techniques of saxophone playing*. Bärenreiter, Kassel, 2010.

[37] M. Young. Nn music: improvising with a 'living'computer. In *International Conference on Computer Music Modeling and Retrieval*, pages 337–350, 2008.