Proceedings of the 10th International Conference on Arts and Humanities, Vol. 10, Issue 1, 2023, pp. 20-37

ISSN 2357 – 2744 online
DOI: https://doi.org/10.17501/23572744.2023.10102

eTu{d,b}e: DEVELOPING AND PERFORMING SPATIALIZATION MODELS FOR IMPROVISING MUSICAL AGENTS

Pocius K*, Davis T and Cusson V

CIRMMT/IDMIL, Schulich School of Music, McGill University, Canada

Abstract: The eTube is a simple acoustic instrument outfitted with a microphone and a two-button controller which we program to facilitate interaction between an improviser and improvising musical agents. Through an iterative research-creation process, we develop and perform various etudes with the eTube and musical agents, generating new knowledge through musical creation. Through this process, we have developed two interactive spatialization systems for the eTube.

We begin by describing the *eTu{d,b}e* framework which refers to the eTube instrument and a series of improvised etudes based on human-computer musical interactions. An overview of the instrument is presented, as well as the existing systems as they were when Pocius began the spatialization portion of this project. Secondly, we outline two interactive spatialization systems designed for improvised performance with the eTube and musical agents. An overview of the updates made to the *eTu{d,b}e* framework is presented, followed by a description of the two spatialization systems. Finally, six different performances are presented as case studies to emphasize the advancements and challenges of the project as new approaches are developed.

Keywords: musical agents, augmented instrument, improvised performance, spatial audio

Introduction

eTu{d,b}e is a performance framework where improvising "musical agents" (MA) and improvisers collaborate in musical performance wherein we ask the question, how can spatializing software facilitate interaction between MAs and improvisers (Tatar and Pasquier 2019)? The eTube is an augmented instrument that combines a woodwind mouthpiece, a polyvinyl chloride (PVC) tube, and a custom controller interface. Inspired by Bowers and Archer's "infra-instrument" (Bowers and Archer 2005), the eTube's sound production is purposefully limited, yet it proposes an intriguing sound world utilizing idiomatic contemporary woodwind techniques. Augmented with a simple two-button controller, the eTube enables communication with the MAs, influencing listening and interaction settings in the software. We use a research-creation process which includes methods and theories from both academic and creative domains and is undertaken by an interdisciplinary team that generates knowledge through the act of creation. Our research-creation process involves critically listening to recordings of rehearsals and performances which then inform new software developments like controller mappings and updated corpora recordings.

*Corresponding Author' s Email: edrina@uitm.edu.my

The Augmented Instrument

Augmented musical instruments (AMIs) are acoustic instruments modified through additional sensors while maintaining their original acoustic sound production capabilities (Miranda and Wanderley 2006, 21). The acoustic sound-producing parts of the AMIs are often used to create sonic material to then be manipulated by synthesis systems. These systems can be controlled by mappings between the sensors and the synthesis systems (Meneses 2022, 1). These mappings can allow for a single interface to generate various types of sounds based on their mappings to parameters of various programs (Hunt, Wanderley, and Paradis 2003). While this allows the performer to directly control the sound synthesis and/or processing performed while maintaining their existing expertise in instrumental technique and familiarity with the performance interface, it provides many of its own unique challenges (Meneses 2022, 15).

Additionally, as highlighted by Cook (2001), AMI designers must also consider what is possible for a performer to do without affecting the acoustic sound production of their instruments. Cook refers to possible gestures as “spare bandwidth” and illustrates this concept by investigating trumpet players. The trumpet requires performers to use three fingers on their right hand to engage their valves, leaving the “spare bandwidth” of the remaining fingers and the left hand available to perform on the added sensors.

Adding spatialization to these systems further complexifies these challenges, adding issues of monitoring for the performer as outlined by Berweck (2012) and Nixdorf and Gerhard (2006), and increasing the need for performers to properly tune not only their instruments and sensors but the loudspeaker systems as well (Tremblay 2017), so that they may properly blend with or contrast against the electronic material. As “spare bandwidth” needs to be considered, more complex mappings to control multiple sound sources or more general parameters of these spatialization algorithms may need to be implemented to result in spatialization systems that are pleasingly complex for a listener while also being simple enough that they do not obstruct the creative expression of the performer.

The eTube

The eTube is an AMI constructed from a 2.54 cm diameter, 219 cm long cylindrical polyvinyl chloride (PVC) tube with a simple two-button controller and baritone saxophone mouthpiece (Davis et al. 2023, 1). Conceived of and designed by Vincent Cusson in 2021 in collaboration with Davis, the eTube is inspired by the notion of the “infra-instrument,” which is a limited or simplified instrument that is augmented with electronics (Bowers and Archer 2005). According to Bowers and Archer, since infra-instruments produce simple timbres and have limited technical facility they “can work very well with live processing... [since] more spectral-temporal latitude is available for augmentation” (Bowers and Archer 2005).

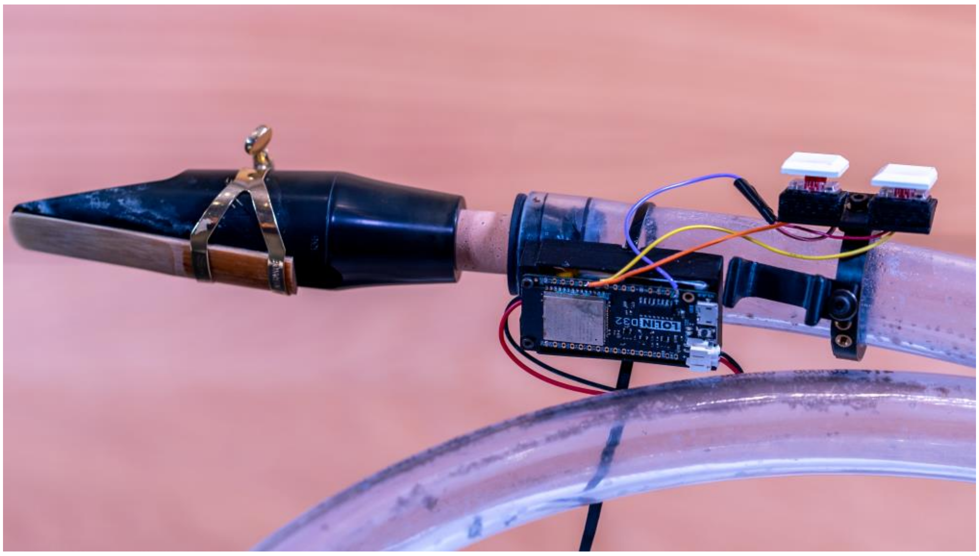

Figure 1: The eTube instrument.

Figure 2: From left to right, Greg Bruce, Tommy Davis and Maryse Legault performing the eTube.

The eTube controller was designed by Cusson in collaboration with Davis and uses available low-cost electronics including 3D-printed parts to anchor the controller to the tube[1]. The design is inspired by the saxophone's keys and right-hand thumb rest (Davis *et al.* 2023, 1). The controller connects wirelessly to a computer via Bluetooth.

The original objective was to design a simple controller that would promote interaction between Davis and the improvising musical agents (MA) in performance (Tatar and Pasquier 2019). The controller in effect transfers certain decisions that were only adjustable by Cusson via the laptop during performances and allows Davis to manipulate these settings in real-time. Both buttons are programmed with three different gestures: one click, a double click, and a long click. Combining these three gestures allows us nine different inputs from the controller that may be programmed to adjust settings in the MAs.

The simple controller helps guide the design process and imposes certain limits on the mappings used to communicate with the MAs. Early in the process, the group would change the mappings for

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

different performances depending on the corpora, collaborators, performance context (Bown *et al.* 2013), and the MA(s) used. However, these mappings remain consistent in the case studies below. Maintaining consistent mappings for the MAs helped Davis to develop familiarity with the performance interface, allowing him to "reach higher levels of subtlety" (Tremblay, Boucher, and Pohu 2007, 3) during performance.

Our setup includes a microphone to transmit the eTube sound to the MAs via an audio interface. We use an Electro-Voice RE20 microphone, which is often used for radio broadcasts or as a kick drum microphone.[2] It has a heavy-duty internal pop filter, allowing it to handle the eTube's direct air pressure, and its cardioid pick-up pattern helps to reject speaker noise without feeding back. Since the eTube performance practice integrates onstage movement, we needed a microphone that wasn't prohibitively expensive, but that was robust enough in case it is bumped or splashed with condensation built up in the eTube. It is a dynamic microphone, which helps capture the wide dynamic range of eTube performance techniques such as slap tonguing.

The eTu{d,b}e Research-Creation Framework

Research-Creation is a relatively recent interdisciplinary approach that combines academic research with creative practices. It is often associated with fields such as the arts, design, and media studies. The methodology emphasizes the integration of research and creative expression, blurring the traditional boundaries between the two (Østern et al. 2023). This research method aims to generate knowledge through the act of creation, with the creative process itself becoming a mode of inquiry. Our work 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. 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. Accordingly, we want to emphasize the importance of effectively communicating and disseminating findings as public performance or outreach, in addition to academic discourse. While institutional recognition of research-creation remains a challenge, our project stands out as a fitting example of the need for this methodology.

Through an iterative process of improvising, listening to concert recordings, and designing hardware the team is continually learning about the MAs through testing and performance (Rowe 2004, 7). Performances are both one of our research outputs, and our methodology for exploring interaction with the MAs and the eTube. This helps us display specific solutions for interacting with MAs, which inform the next developments in the project through a self-reflective process.

eTu{d,b}e framework was first developed by Cusson and Davis in 2021 (Cusson and Davis 2022). It adapts existing MAs in a flexible performance framework as described in Section 2.3. The title eTu{d,b}e refers to the name of the eTube and to a series of improvised etudes based on human-computer musical interactions (Cusson and Davis 2022). The curly brackets are borrowed from computer terminology. This syntax is used to indicate that both letters are interchangeable, representing the intertwined relations between the eTube and etudes on human-computer interaction.

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

The curly brackets also contain the acronym for decibels (dB), which measures the relative loudness of sounds.

Through eTu{d,b}e performances, the group explores the interactive elements of pre-existing MA systems, such as DYCI2 (Nika et al. 2017), Construction III (Kafejian 2023) and Spire Muse (Thelle and Pasquier 2021). This exploration initially included agent listening settings, corpus files and analysis, eTube controller mappings, and number of agents. With the help of Pocius, new spatialization models have been added to expand the artistic and research tools and possibilities.

When Pocius joined the project in 2022, the performances used DYCI2 (Nika et al. 2017) and Construction III (Kafejian 2023). Button mappings for Spire Muse (Thelle and Pasquier 2021) had been completed by Cusson and had been used during Davis' residency at Simon Fraser University's (SFU) Metacreation Lab, but had not been used in performance. Further information about the agents and the controller mappings can be found in our 2023 NIME paper (Davis et al. 2023).

As previously mentioned, this project has been heavily inspired by the infra-instrument concept described by Bowers and Archer (Bowers and Archer 2005). As a result, most aspects of this project have been approached with similar limitations in mind. Corpora curation has mostly been limited to eTube and saxophone sounds and almost all of the corpora presented in the case studies have been recorded by Davis (see Section 4). The team could have used numerous recordings of other sounds for the corpora but wished to keep the agents as an extension of Davis' ongoing exploration of the eTube's sonic identity.

As mentioned in (Davis *et al.* 2023)

> "we were familiar with the improvised material in the MA's corpora because we had recorded it ourselves, and this helped us to evaluate the MA's listening settings. Using corpora that were homogenous [sic] 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 the performance of an instrument with MAs as two separate realities."

The one exception to this is the modular synthesizer corpus which Pocius designed for this project (further explored in Section 4.5), which represents a new direction for the project. Even with this new corpus, we continue to find the eTube's limited acoustic capabilities, paired with the fact that Pocius recorded the corpus and is, therefore, familiar with all its contents, to be enlightening in terms of how these agents analyze and respond to our inputs. The analogue synthesizer recordings are also much fuller with less dynamic contrast than the eTube corpora. This created a contrasting sonic environment compared to the eTube corpora used in the same performance. In addition, the complex electronic sounds provided a contrasting soundscape that compliments Davis' noisy contemporary techniques on the eTube and inspired new sonic combinations.

Spatialization systems such as those used in Stevie Wishart's The Sound of Gesture (Brummer et al. 2014) or Alyssa Aska's Fayum Fragments (Aska and Ritter 2016), utilize additional sensors for their spatialization. Our spatialization system uses audio from the onstage microphone for feature extraction and avoids additional controller mappings or added sensors to control meta-parameters within the spatialization algorithm (Section 3).

Using feature extraction keeps the existing instrument simple, allowing the eTube performer to control the spatialization using their existing acoustic vocabulary while directing their "spare bandwidth" (Cook 2001) to interact with the controller and engage in musical conversation with the MAs.

Musical Agent Descriptions

Three existing MAs are used with the eTu{d,b}e framework and this includes the Creative Dynamics of Improvised Interaction (DYCI2) by Jérôme Nika and colleagues, Spire Muse by Notto Thelle and colleagues, and Construction Tools for Interactive Performance (CTIP) by Sergio Kafejian (Nika et al. 2017; Thelle and Pasquier 2021; Kafejian 2023). A new version of CTIP has been updated by Cusson and is called Construction III.

Deeper descriptions of these agents and our previous performance practice can be found in our 2023 NIME paper (Davis et al. 2023, 3). Updated versions of the DYCI2 and Spire Muse software are named dicy2 and Co-Creative Communication Platform (CCCP) respectively, although we do not address the most recent updates below (Nika et al. 2022; N. J. Thelle and Wærstad 2023).

We will briefly describe each MA's pre-existing spatialization functions.

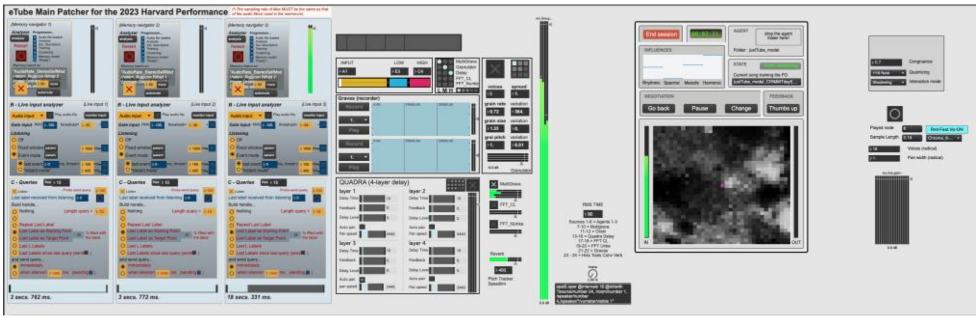

Figure 3: The main interface for the performance Max patch containing all three agent softwares. Three DYCI2 agents are seen on the left-hand side, showing the interface for the listening and rendering parts of the agent. Construction III is placed in the middle, showing settings and levels for each effect and input level at the top. The master gain for the patch and the interpolation for the spatialization is placed in beside the interface for Construction III. The interface on the right-hand side shows the interactive settings for Spire Muse, including input and output levels for the agent, the large square displaying the self-organized map of samples and current query, and buttons and sliders to manipulate the playing and listening modes.

DYCI2

DYCI2 performances often uses multiple DYCI2 agents, each with separate corpora and listening settings (Nika et al. 2017). As DYCI2 agents use stereo audio files for their corpora, the agents will maintain whatever stereo imaging is already captured in the original corpora.

As each DYCI2 MA pulls from a single audio file at a time, their musical vocabulary is quite limited. This has led the eTube team to explore performances using multiple agents which are often used alongside a reactive system such as Construction III.

Spire Muse

Spire Muse employs one large corpus, trained offline using a patch based on the Musical Agent Based on Self-Organizing Maps (MASOM) architecture (Tatar and Pasquier 2017). The Spire Muse agent patch makes use of Max's polyphonic function to build its musical output, layering up to 16 voices to create new musical phrases. In the initial patch, each of these voices functions in mono, and a random panner on the synth voice is triggered when a new musical fragment is called for that voice. This creates a varying stereo image throughout a performance, but these results have little relation with the inputs or outputs used to generate them.

Construction III

The Construction III system uses pitch analysis to route a monophonic input into various effects processing modules. This MA has some built-in spatialization module via a matrix mixer, with outputs in stereo, quadraphonic, or octophonic (Kafejian 2023).

The New Spatialization Systems

Pocius has implemented two new spatialization systems for this project. One system builds on the performance patch using DYCI2 and Construction III. This system uses a series of spatialization presets which can be manipulated by both the eTube performer and the computer performer. The second system uses Spire Muse and implements the 16 channels used for the sound generation of the agent as a cluster of objects which can be controlled by various parameters derived from the pitch chroma analysis of the eTube's sound. More technical information can be found in Pocius' master's thesis (Pocius 2023, 37–54)

Preset System - DYCI2 and Construction III

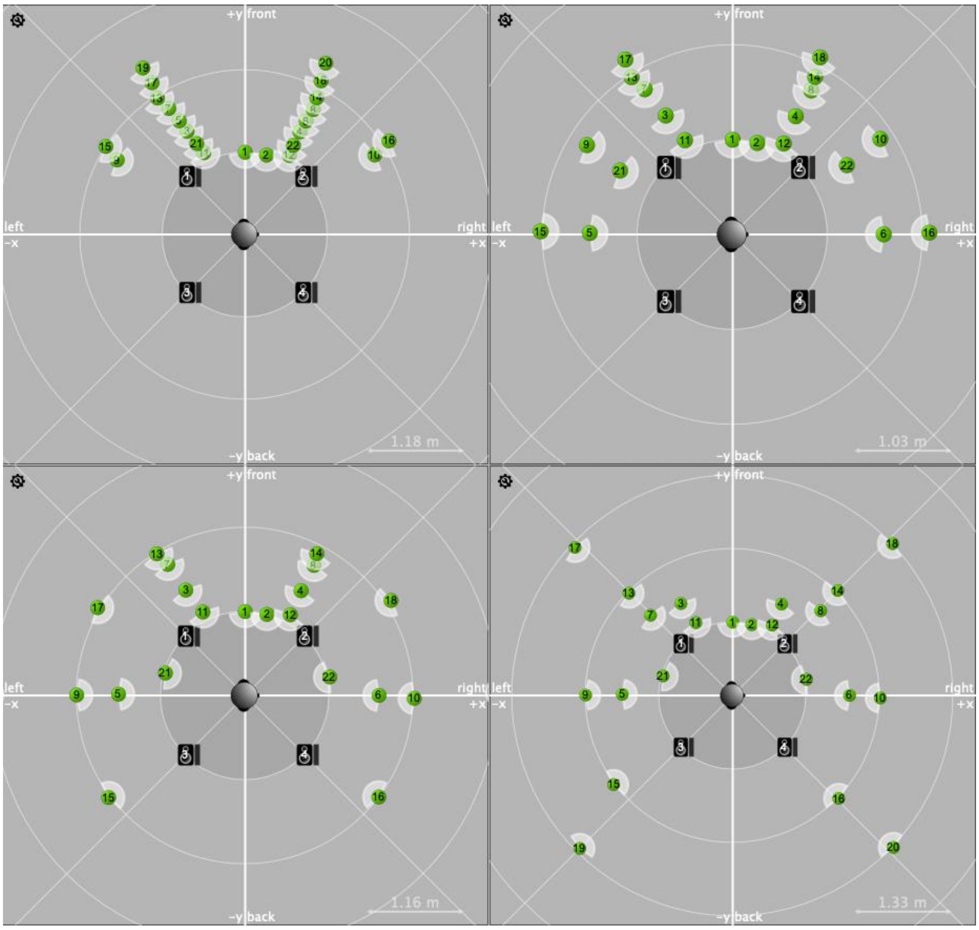

Figure 4: Four of the presets used for the preset spatialization system. The green circles represent the sound sources.

As discussed in sections 2.3.1 and 2.3.3, the patch using DYCI2 and Construction III already had some limited spatialization in place, as each DYCI2 agent uses the spatial vocabulary embedded within the stereo sound file used for its corpus, while Construction III contained several sets of stereo and quadraphonic effect patches. Cusson had implemented a matrix mixer combining both patches to map the outputs to the final system output channels, which allowed for flexible mappings of channels during performance by the computer performer but could not interpolate between settings smoothly. Considering the existing controller mappings used to interact with the agents, it was deemed impractical for the eTube performer to control the matrix via the controller during performance.

Pocius broke out these outputs into 24 independent sources which are then fed into Spat5 (Carpentier, Noisternig, and Warusfel 2015). Focusing on Bowers and Archer's infra-instrument principles, Pocius wanted to create a system that could have a relatively simple interaction which would not require too much "spare bandwidth" to use but could still result in a complex spatial result. Instead of controlling each source individually, Pocius programmed four presets in Spat5, ranging from all sources placed in front to the sources surrounding the audience (Figure 4). These presets focus on placing the individual agents' sources closer to the loudspeakers so that they can be more clearly distinguished, with the processing placed in various locations between loudspeakers. Inspired by the envelope followers used in the works by Graham et al. and Graham and Bridges (2017; 2014), the eTube performer can interpolate between adjacent presets via an envelope follower on the eTube microphone, while the offset and modulation depth is controlled by the computer performer using an AKAI MIDImix controller, allowing them more control over the overall spatial shape of the performance as needed. The placement and behaviour of the virtual sound sources will respond based on the amplitude of Davis' playing as captured by the mic, and Pocius can control how quickly these sources respond and how quickly they change between preset locations. This flexibility is desirable so that performers can manipulate these presets to suit individual moments during an improvisation, without the need to pre-program trajectories or other time-sensitive material which is undesirable in an improvisational context, particularly one using MAs whose outputs remain at least partially unpredictable.

The Cluster System - Spire Muse

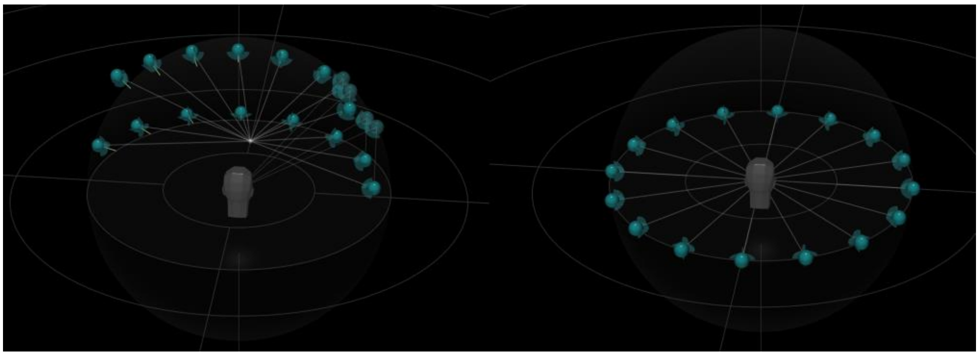

Figure 5: Two of the clusters used for the cluster spatialization system. The blue balls represent sound sources.

As discussed in section 2.3.2, Spire Muse originally had a random panner, with each of the 16 mono voices of the synthesis engine being randomly panned with each new grain played. Pocius deconstructed this so that each of the 16 voices was sent to its own channel. 16-channel clusters were then designed in Spat Revolution (Figure 5), a commercial spatialization software. Inspired by the harmonic analysis patches used to control flocking algorithms outlined by Graham et al. (2014) the rotation of this cluster along the Z axis is controlled by the pitch chroma of the eTube's signal, captured via a Fluid Corpus Manipulation (FluCoMa) (Tremblay, Roma, and Green 2021) pitch tracker. The confidence of the pitch tracker is used to control the scale of this cluster, allowing the eTube performer to control the distribution of the cluster throughout the speaker system. Various extended techniques with fluctuating breath noise will affect the confidence of the pitch-tracking algorithm.

The Spire Muse cluster's shape is preserved as it scales and rotates, allowing the 16 outputs to be more easily separated while maintaining coherence as a single entity. The eTube's overall pitch contour, as well as inherent micro-gestures from techniques such as swinging the eTube, affect the rotation and scaling of the cluster. This system has been updated with the latest version of Spire Muse created by Thelle for the Co-Creative Spaces project (Thelle and Wærstad 2023), but has only recently seen concert use with our project.

Performance Case Studies

Fall 2022 Workshop at CIRMMT

Figure 6: The microphone setup used for the workshop.

In fall 2022, Pocius and Davis tested the spatialized Spire Muse system described in section 3.2 and recorded material for the fixed media work eTu{d,b}e de labo #1 by Kasey Pocius at the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT).

This setup used the "all tube" Spire Muse corpus, which consists only of recordings of various short improvisations made in 2022 on the tube. The controller mappings for Spire Muse made by Cusson and Davis were implemented.

Spat Revolution's binaural monitoring system was used to allow Davis and Pocius to hear the spatialization of the agents without the need for loudspeakers. The agent was spatialized as a cluster consisting of two semi-circles of eight sources placed facing the eTube performer. The RE20 that was used to engage the agents and the 16-channel output of Spire Muse was captured in a Digital Audio Workstation (DAW), along with the binaural rendering of the spatialized MA, while a microphone array recorded a stereo image of the eTube performance from various proximities. This session was run at a sampling rate of 96 kHz to capture additional harmonics from the eTube for future processing, and while this increased computational load significantly, no technical glitches were discovered in either the MA or the spatialization system during this recording.

Codes d'accès Performance

The Codes d'accès performance was the first public performance using Pocius' system outlined in 3.1, with the MIDI controller faders additionally being used to control the agent's levels and the processing while Davis continued to have control over the agents and processing via the eTube controller.3 Overall audience feedback was good, and this performance suggests that the spatialization system required little "spare bandwidth" and did not affect Davis' ability to play with the MAs, while still producing an artistically satisfying result for all those involved.

---

[3] https://codesdacces.org/evenement/coda-2022-echanges/

The water tube recording from the NIME 2022 performance (Cusson and Davis 2022) was used for DYCI2's agent one. Agent two consisted of a short low drone recorded with a Barcus Berry 4000XL contact microphone.4 This recording was selected for its timbral qualities, which were well captured by the contact microphone on the tube. Agent three included articulated eTube harmonics, which added rhythmic variation and response.

Pocius performed with their MIDI controller near the front of the stage and contributed additional mixing and processing. They aimed to shape the overall improvisation via mixing, adding a focus on the performative approaches of both the computer and acoustic performers. This was in contrast to past performances which used a more hands-off approach to the performance allowing Davis more space for interaction with the agents via the controller.

Figure 7: Davis performing Enfants, apprenez-nous à parler in the MMR. Note the microphone setup used.

Enfants, apprenez-nous à parler

Premiered as part of a live@CIRMMT concert in 2022, 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

---

[4] https://kmcmusic.com/

qualities of the eTube.5 The work developed a notation system for the eTube, and features composed and improvised sections using spatialized fixed audio, real-time processing, and DYCI2 agents. Pocius helped expand the spatialization done in Spat5. The eTube controller was used to trigger fixed cues in the Max patch. Programmed by Lauvray, these cues trigger several 16-channel fixed audio files with one ring of eight at ear level, and another ring of eight placed above the listener. These cues also manipulate the routing of effects sent to Spat5 via a matrix mixer, as well as manipulate the trajectories for these effects and the MAs. While some cues contain fixed trajectories for the processing, many use random placement of sources. Spat5 is used to spatialize the processing and the MAs in a 16-channel setup similar to the setup used for the audio files.

Pocius designed a final mixing patch in Spat5 allowing them to down-mix or up-mix the 16-channel master to an arbitrary number of outputs. This improves transportability and adaptability in different performance spaces, which was evident during the premiere at CIRMMT's Multimedia Room (MMR).

As described by Davis et al., "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 spatialized trajectories" (Davis et al. 2023, 6).

While this microphone array has largely been successful in capturing spatial gestures of the eTube, it can often be difficult to secure three matching dynamic microphones, even in academic institutions. As described in section 4.6, the input patch has been updated by Lauvray to support a stereo mic pair, which is then used to derive a third channel with content common to both microphones.

Tommy Davis' Doctoral Lecture-Recital

Davis' lecture-recital provided an overview of the eTu{d,b}e project, with one improvisation with the spatialized DYCI2 and Construction III system, an improvisation with the spatialized Spire Muse system and a performance of Enfants, apprenez-nous à parler.6 The lecture-recital gave a project overview and each improvisation aimed to showcase the differences between the MAs. This lecture-recital was given twice in two weeks on two separate quadrophonic systems, a small loudspeaker setup provided to Davis by the McGill Digital Composition Studio (DCS) performed in a classroom, and a Public Address (PA) system in Tanna Schulich Hall provided by the DCS.

The improvisations with DYCI2 and Construction III used the same corpora as the Codes d'accès performance (Section 4.2), and the Spire Muse improvisation used the "all tube" corpus, but this time the cluster was configured as 16 sources at ear level, as there were no height channels present.

There was little time to troubleshoot at each location. However, the eTube's Bluetooth connection provided no issues, even with the computer located at the back of the hall and the eTube performer onstage, and the virtual loudspeaker setup used to down-mix Enfants, apprenez-nous à parler worked

---

[5] https://www.cirmmt.org/en/events/live-cirmmt/improvising-new-winds

[6] https://www.mcgill.ca/music/channels/event/doctoral-lecture-recital-tommy-davis-saxophone-343602

with no further adjustments necessary. These performances confirmed to the group that multiple agents could be used in one concert performance provided there was enough time to change patches and train DYCI2 MAs.

The quadraphonic system was installed with two speakers placed onstage flanking the performer and two speakers at the back of the hall behind the audience. Tanna Schulich Hall has raked seating and the back of the hall is higher than the stage, and Davis had trouble hearing the back speakers. Lauvray's piece included an effect which moved clockwise through the quadraphonic system and required Davis to trigger the effect in rhythmic succession, which was challenging since the back speakers were hard to hear onstage. For the improvisation with Spire Muse, Davis decided to play sparsely to allow the MA more space to contribute. Spire Muse's output was also very sparse, and although Davis was not convinced of the resultant musical result in the moment, he maintained the same sparseness since this was a demonstration in a lecture-recital context. Since this performance, we have had success performing sparse material with Spire Muse.

2023 North American Saxophone Alliance Biennial

Following the success of Tommy Davis' lecture-recital (Section 4.4), Davis and Pocius decided to combine both spatialized systems into one patch for their performance at the 2023 North American Saxophone Alliance (NASA) Biennial Conference, hosted at the University of Southern Mississippi.7 In an attempt to help the audience separate Spire Muse and DYCI2, Pocius created a Spire Muse corpus by curating several multi-tracked modular synthesizer recordings, while DYCI2 would use the corpora used in the previous performances (outlined in Section 4.2). While the group requested a quadraphonic system from the conference, they were only able to supply the group with a stereo PA system. During sound check, Pocius used Reaper to mix the back channels into the front channels and selected the cluster outlined in section 4.1. Davis and Pocius used pink noise bursts to balance the levels and spectrum of both speakers during their sound check, followed by a short rehearsal of the improvisation where we found the spatialization to still provide satisfying results despite the lack of a full surround setup.

The first half of the performance used DYCI2 and Construction III. Near the end of the first half, DYCI2 was outputting more soloistic material and Davis decided to take a supportive role, contextualizing the MA's output as a solo by performing rhythmic slap tongues alongside the agent. This was an exciting first opportunity to perform with Pocius' analogue synthesizer corpus which was featured in the second half with Spire Muse. Spire Muse was able to closely match Davis' high overtones and textural teeth-on-reed techniques at the beginning of the second half, resulting in a complimentary dialogue between Davis' acoustic textures with the synthesizer corpus.

[7] https://www.saxophonealliance.org/conference.asp

Performance at IIICON2023

Figure 8: Davis and Pocius performing at Holden Chapel during IIICON2023.

Davis and Pocius decided to use both systems for one improvisation followed by a performance of Enfants, apprenez-nous à parler for the Instruments, Interfaces, Infrastructures: Interdisciplinary Conference on Musical Media 2023 (IIICON2023) hosted at Harvard University.8 Throughout rehearsals at the DCS for IIICON2023, Davis and Pocius found the "all tube" Spire Muse corpus resulted in responses with a more varied dynamic range than the modular synthesizer corpus. Additionally, we found that reducing the segment size used by DYCI2 for analysis would result in a more varied sonic output, with phrases that would be impossible to play on the acoustic tube, helping to better separate it from the live performer. As this performance was presented in quadraphonics, the circular cluster used for Davis' lecture-recital (Section 4.4) was selected. To better accommodate the setup at the conference, this performance instead used a stereo pair of AKG C414 XLS microphones, set to cardioid. For Enfants, apprenez-nous à parler, the input patch was modified by Lauvray so that an artificial center channel could be created by mixing the signal from the two microphones together, and no other modification to the electronics would be needed. While levels from sound check were not maintained during the performance, this concert is largely deemed a success by the group. Additionally, the time needed to switch between the patches for the improvisation and for Enfants, apprenez-nous à parler was sufficient to give the audience information about the eTube and the eTu{d,b}e framework, helping the audience to better contextualize the MAs and the spatialization systems.

The C414 microphones caused the agents to react compared to the RE20s. Although Davis had practiced with a two-C414 setup, the agent's reactions in that venue were surprising. The agent's threshold seemed to trigger much more easily with the C414s, and as a result, Davis consciously played quieter and kept the eTube approximately 30 cm from the C414s, never placing the eTube directly on the microphone grill as he often would with the RE20. The Doppler effect was much more pronounced with the C414s as Davis rotated the eTube back and forth in front of the C414s at

[8] https://sites.harvard.edu/instruments-interfaces-infrastructures/

different rates. Davis played with this rhythmic aspect and Pocius responded later by playing back recordings of this gesture using Construction III.

Limitations and Future Work

Monitoring

Monitoring for the eTube performer remains an issue, similar to those outlined by Berweck (2012) and Nixdorf and Gerhard (2006). Most halls have a central mixing point which is ideal for hearing all sonic events. The performers are often located close to the front speakers, far from the ideal listening position, and struggle to hear sounds that are diffused in other areas of the space. In addition, bone conduction transfers sound from the woodwind mouthpiece to the inner ear via the teeth and skull and vocalizations are also passed through these structures (Braasch 2019, 35-43). 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.

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 (Federman and Ricketts 2008). However, bone conduction for wind instruments is increased when the ear is occluded, so in-ear monitors may mask softer electronic sounds, especially if the monitors are set to a lower volume for aural health (Albrecht et al. 2017). While binaural headphone monitoring has proved to help during studio recordings, this added technical overhead for performance is an unwanted complication for the project at this time. It has proven useful in recording environments where we have the necessary time for setup (Bauer et al. 2022). The in-ear monitors would provide a clearer and more separated sound for the performer, but the hall's acoustic reflections and interactions are not respected in binaural mixes. This limits an improviser's ability to engage with live features in the hall, diminishing the real-time connection between performers and the space (Bauer et al. 2022). Since directionality, movement, and interaction with acoustic space are important aspects of Davis' eTube practice, this has influenced our choice to limit in-ear monitoring in live performances, in light of some benefits mentioned above.

If the speakers are placed correctly, spatialization in the back speakers can improve the eTube performer's ability to hear the agents, as agents in the back speakers will be heard more directly by the eTube performer than agents placed in the front pair, as the front pair often faces away from the performer. This is, however, dependent on the acoustics of the hall.

Conclusion

We presented two spatialization models developed by Kasey Pocius which facilitate interaction in improvised performance between an improviser and MAs. These models were adapted for use with three existing musical agents as part of the eTu{d,b}e framework. One model used preset interpolation and an envelope follower and the second model used a pitch tracker to control the scale and rotation of a cluster of objects. Maintaining limitations inspired by the infra-instrument concept, the existing eTube setup was utilized. Six case studies explore our interdisciplinary research-creation process and artistic output, highlighting the successes and challenges of implementing these technologies in improvisations with MAs.

Acknowledgements

The eTube has been developed thanks to Student Research Awards (2021-22, 2022-23, 2023-24) and Inter-centre 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 Marie-Chantal Leclair, Robert Hasegawa, Guillaume Bourgogne, and Marcelo M. Wanderley for their support and guidance during this research. Thank you to Jérome Nika, Sergio Kafejian, Notto Thelle, Ben Carey and Henning Berg 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. Thanks to Yves Méthot at CIRMMT and to Richard McKenzie at the DCS at McGill for their support throughout the project

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

Declaration of Interest Statement

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

References

Albrecht, Robert, Jussi Jaatinen, and Tapio Lokki (2017). Electronic hearing protection for musicians. In Proceedings Sound and Music Computing Conference, 306–13.

Aska, Alyssa, and Martin Ritter. (2016). Approaches to real time ambisonic spatialization and sound diffusion using motion capture. In Proceedings of the International Computer Music Conference.

Bauer, Valentin, Dimitri Soudoplatoff, Leonard Menon, and Amandine Pras. (2022). 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.

Bowers, John, and Phil Archer. (2005). Not hyper, not meta, not cyber but infra-instruments. In Proceedings of the International Conference on New Interfaces for Musical Expression, 5–10.

Bown, Oliver, Arne Eigenfeldt, Aengus Martin, Benjamin Carey, and Philippe Pasquier. (2013). The Musical Metacreation weekend: Challenges arising from the live presentation of musically Metacreative systems. In Proceedings of the Artificial Intelligence and Interactive Digital Entertainment Conference, 27–34.

Braasch, Jonas. (2019). Hyper-specializing in saxophone using acoustical insight and Deep Listening skills. Edited by Rolf Bader, Marc Leman, and Rolf-Inge Godoy. Vol. 6. Current Research in Systematic Musicology. Cham, Switzerland: Springer.

Brummer, Ludger, Gotz Dipper, David Wagner, Holger Stenschke, and Jochen Arne Otto. (2014). New developments for spatial music in the context of the ZKM Klangdom: A Review of technologies and recent productions. Divergence Press 3 (1).

Carpentier, Thibaut, Markus Noisternig, and Olivier Warusfel. (2015). Twenty years of Ircam Spat: Looking back, looking forward. In Proceedings of the International Computer Music Conference, 270–77.

Cook, Perry. (2001). Principles for designing computer music controllers. In Proceedings of the International Conference on New Interfaces for Musical Expression, 1–4.

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

Davis, Tommy, Kasey Pocius, Vincent Cusson, Philippe Pasquier, and Marcelo M. Wanderley. (2023). eTu{d,b}e: Case studies in playing with musical agents. In Proceedings of the International Conference on New Interfaces for Musical Expression.

Federman, Jeremy, and Todd Ricketts. (2008). Preferred and minimum acceptable listening levels for musicians while using floor and in-ear monitors. *Journal of Speech, Language, and Hearing Research* 51 (1): 147–59.

Hunt, Andy, Marcelo M Wanderley, and Matthew Paradis. 2003. The importance of parameter mapping in electronic instrument design. *Journal of New Music Research* 32 (4): 429–40.

Kafejian, Sergio. (2023). The use of an interactive system in different improvisational contexts. Pre-print, received via personal communication with Tommy Davis

Meneses, Eduardo. (2022). Iterative design in DMIs and AMIs: Expanding and embedding a high-level gesture vocabulary for T-Stick and GuitarAMI. PhD thesis, McGill University (Canada).

Miranda, Eduardo Reck, and Marcelo M Wanderley. (2006). New Digital Musical Instruments: Control and interaction beyond the keyboard. Vol. 21. AR Editions, Inc.

Nika, Jérôme, Ken Déguernel, Axel ChemlaRomeu-Santos, Emmanuel Vincent, and Gérard Assayag. (2017). DYCI2 agents: Merging the 'free,' 'reactive,' and 'scenario-based' music generation paradigms. In Proceedings of the International Computer Music Conference.

Nika, Jérôme, Augustin Muller, Joakim Borg, Gérard Assayag, and Matthew Ostrowski. (2022). Dicy2 for Max. Ircam UMR STMS 9912. https://hal.science/hal-03892611.

Østern, Tone Pernille, Sofia Jusslin, Kristian Nødtvedt Knudsen, Pauliina Maapalo, and Ingrid Bjørkøy. (2023). A performative paradigm for post-qualitative inquiry. *Qualitative Research* 23 (2): 272–89.

Pocius, Kasey. (2023). Expanding spatialization tools for various DMIs. Master's thesis, McGill University.

Rowe, Robert. (2004). Machine musicianship. MIT Press.

Tatar, Kivanc, and Philippe Pasquier. (2017). MASOM: A musical agent architecture based on Self-Organizing Maps, affective computing, and variable Markov models. In Proceedings of the International Workshop on Musical Metacreation, 1–8.

Tatar, Kivanç, and Philippe Pasquier. (2019). Musical agents: A typology and state of the art towards Musical Metacreation. *Journal of New Music Research* 48 (1): 56–105.

Thelle, Notto J. W., and Philippe Pasquier. (2021). Spire Muse: A virtual musical partner for creative brainstorming. In Proceedings of the International Conference on New Interfaces for Musical Expression.

Thelle, Notto J. W., and Bernt Isak Wærstad. (2023). Co-Creative Spaces: The machine as a collaborator. In Proceedings of the International Conference on New Interfaces for Musical Expression.

Tremblay, Pierre Alexandre. (2017). Tuning to trust: System calibration as creative enabler. In Proceedings of the International Computer Music Conference, 179–84.

Tremblay, Pierre Alexandre, Nicolas Boucher, and Sylvain Pohu. (2007). Real-time processing on the road: A guided tour of [Iks]'s Abstr/Cncr Setup. In Proceedings of the International Computer Music Conference.

Tremblay, Pierre Alexandre, Gerard Roma, and Owen Green. (2021). Enabling programmatic data mining as musicking: The Fluid Corpus Manipulation Toolkit. *Computer Music Journal* 45 (2): 9–23.