ACMC2015 - *MAKE!*

Proceedings of the Annual Conference of the Australasian Computer Music Association

Hosted by the Faculty of Arts and Social Sciences
University of Technology Sydney

Wednesday 18th- Saturday 21st November 2015

Proceedings of the 2015 Conference of the Australasian Computer Music Association
ISSN 1448-7780

Proceedings of the 2015 Annual Conference of the Australasian Computer Music Association Sydney Australia

18th – 21st November 2015

Keynote Speaker:
Associate Professor Cat Hope

Published by:
The Australasian Computer Music Association

http://acma.asn.au
November 2015

ISSN 1448-7780

Proceedings edited by Ian Stevenson

All correspondence with authors should be sent directly to the authors.

General correspondence for ACMA should be sent to info@acma.asn.au

The peer review process applied to full papers published in these proceedings involved impartial and independent assessment and review of the papers before publication, conducted by independent, qualified experts as required by section 6.5 of the 2015 Higher Education Research Data Collection Specifications of the Australian Government Department of Education and Training.

TABLE OF CONTENTS

FULL PAPERS

Ros Bandt 5
CREATING SOUNDING SPACES- THE WHY? FACTOR

Rodney Berry, Ernest Edmonds, Andrew Johnston 13
UNFINISHED BUSINESS- SOME REFLECTIONS ON ADOPTING A PRACTICE-BASED APPROACH TO TECHNOLOGICAL RESEARCH AS AN ARTIST

Colin Black 19
EXPLORING A DISEMBODIED VOICE IN SURROUND SOUND MEDIA SPACE

Brigid Bourke 26
ESCAPEE GLOSS-A SYMPHONY OF POLYMEDIA

Benedict Eris Carey, John R. Taylor, Damian Barbeler 34
AN INTERACTIVE, MULTIMODAL INSTALLATION USING REAL-TIME SPECTRAL ANALYSIS

Paul Dunham, Bridget Johnson 40
A FRACTURED EARTH- SONIFICATION AND SPATIALISATION OF SEISMIC DATA WITH MAX:MSP AND SPEAKER MOTION

Tim Gilchrist, Meg Travers, Tara Surry, SKoT McDonald, Morgan Strong, Craig Wale 45
THE SONIC BOLLART

Vincent Giles 50
SCORING FOR GRID CONTROLLERS – PRELIMINARY CONSIDERATIONS OF ACTION, GESTURE, FUNCTIONALITY, AND LIMITATIONS

Nigel Helyer 57
A DIFFERENT ENGINE

Cat Hope, Meg Travers 64
WHO WANTS A TRAUTONIUM MUST BUILD ONE

James Humberstone, John R. Taylor 69
MAKING MUSIC LEARNING FUN- DESIGNING AN INTERACTIVE IBOOK FOR INFORMAL LEARNING

Blake Johnston, Michael Norris 75
CARME- A MULTI-TOUCH CONTROLLER FOR REAL-TIME SPATIALISATION OF GESTURAL BEHAVIOUR

Mary Mainsbridge 81
CRACEA-GESTURAL PERFORMANCE

Charles Martin, Henry Gardner, Ben Swift, Michael Martin 85
MUSIC OF 18 PERFORMANCES- EVALUATING APPS AND AGENTS WITH FREE IMPROVISATION

SKoT McDonald 95
BFDLAC, A FAST LOSSLESS AUDIO COMPRESSION ALGORITHM FOR DRUM SOUNDS OR SIZE ISN'T EVERYTHING

Mark Pedersen, Prue Lang 101
GLACIER MELT- EMBODIED KNOWLEDGE AND THE COMPOSITION OF CHOREOGRAPHIC SOUNDSCAPES

Ian Stevenson 108
PERFORMATIVITY AND INTERACTIVITY-CONCEPTS FOR FRAMING THE PEDAGOGY OF MAKING

Michael Terren, Cat Hope 112
MAP-MAKING TOWARDS AN ONTO-CARTOGRAPHY OF THE DIGITAL AUDIO WORKSTATION

Lindsay Vickery, Talisha Goh 119
MUSIC SCREEN-READING-INDICATIVE RESULTS FROM TWO PILOT STUDIES

Lindsay Vickery 160
THROUGH THE EYE OF THE NEEDLE-COMPOSITIONAL APPLICATIONS FOR VISUAL-SONIC INTERPLAY

ABSTRACTS

Liam Bray and Oliver Bown 135
LUDIC HUMAN COMPUTER CO-CREATION

Ben Byrne 135
SOUND AS MULTIPLICITY IN EXPERIMENTAL MUSIC: LISTENING WITH THOREAU, CAGE AND SERRES

Luke Harrald 135
ACOUSMATIC COMPOSITION IN A BLENDED WORLD

Charles Martin 136
GESTURE STUDY AND CORRESPONDENCES: COMPOSING FOR PERCUSSIVE IPAD APPS

Terumi Narushima, Kraig Grady, Matthew Dabin, Christian Ritz, Stephen Beirne 136
3D PRINTING MICROTONAL FLUTES

Michelle Stead 137
WRITING THE WRONGS OF HISTORY: LISTENING AND THE DISCOURSE OF GENDER IN ELECTROACOUSTIC MUSIC

ARTIST TALKS

Roger Alsop and Brigid Burke 139
GREEN SKIES…. MOMENTS-ARTIST TALK

Andrew Brown 143
IMITATION AND AUTONOMY AS A BASIS FOR MUSICAL AGENCY-ARTIST TALK

Warren Burt 144
“A PLETHORA OF POLYS” – A LIVE ALGORITHMIC MICROTONAL IMPROVISATIONAL COMPOSITION FOR IPAD-ARTIST TALK

Damian Castaldi 149
PLAY ONGAKU-ARTIST TALK

Keith Halpin 151
SONIC ART AS METAPHOR & MYTHOLOGY-ARTIST TALK

STUDIO REPORTS

Cat Hope, Lindsay Vickery, Jean-Michel Maujean 154
WEST AUSTRALIAN ACADEMY OF PERFORMING ARTS, COMPOSITION DEPARTMENT-STUDIO REPORT

Timothy Opie 155
IMPLEMENTING THE NEW BACHELOR OF SOUND PRODUCTION AT BOX HILL INSTITUTE-STUDIO REPORT

Meg Travers and SKoT McDonald 157
THE PERTH ARTIFACTORY-STUDIO REPORT

ACMC2015 – *MAKE!*

REFEREED FULL PAPERS

CREATING SOUNDING SPACES: THE WHY? FACTOR

Dr Ros Bandt
International sound artist composer, performer, sound culture researcher, Director of the on-line Sound Design Project and Gallery, Faculty of Arts
University of Melbourne www.rosbandt.com
www.hearingplaces.com

ABSTRACT

Sound artist and composer Ros Bandt, reflects on "Creating sounding spaces, the why factor?" with reference to the complexities of approaching two recent international commissions, the six channel *Raptor* 2014 for the ListenN festival, USA and a performance piece *Five Faces of Medea* for the German recorder Virtuoso Ulrike Volkhardt with 5 videos, surround sound and live electronics.

1.WHY DO IT? WHY MAKE WORK?

Thinking it through.

In my studio I have a blank canvas. It beckons me with the question mark of empty space for each new work. What do I want to create? For Whom, Where? For how long? Why do I want to do this? A commission? A dream? An invitation? A non–intention; something just growing up by itself? An experiment? Some novel research and development collaboration? A challenge? A computer aberration? The answers to all these questions come together in unique ways which shape the creative outcome. I never sit down at the computer to just make a new piece of computer music because I can. Technical possibilities are available to us, but as artists there has to be a much better reason. There's a tango waiting to happen, people in an acoustic space somewhere.

Two very recent international commissions, *Raptor* and *5 faces of Medea* are case studies for these questions to be answered. They are two very different working models and outcomes. The complex roles of the computers are totally subservient to the overall aesthetic needs of each work, but both works rely heavily on them for the production of the ideas at every stage of the musical process.

2. CASE STUDY I : RAPTOR

Raptor Ros Bandt 15 minutes

5.1 computer composition interpreting an international environmental sound archive for an international festival.

I received an email from Garth Paine director of electronic arts at the university of Arizona to invite me to be one of a number of internationally commissioned works for his ListenN festival in Arizona, 2014. The task was to create a work interpreting the ambisonic sound recordings of the environmental regions he is archiving of the biosphere in Arizona. The sounds recorded by someone else (him or Leah Barclay) had to be the focus of the composition as the whole point of the commissions was to make people aware of the biosphere data base they were creating as a continuing sound laboratory of the natural environment, a worthy task.

Unlike most computer music composers, I have always created my own sound palette and made my own air and hydrophone recordings and rarely used anyone else's samples despite encouragement from The West Deutsche Rundfunk and the ABC to use their sound libraries. An exception is the brutal horse race end of the Bacchylides movement in my Thrausmata *Sonic Archeologies* CD on Move Records. The act of listening and "being in place" is part of the integrity of field composition for me personally. Several days were spent listening to this environmental sound archive. I found the on-line recordings very variable in quality and sonic interest. For the most part as Garth described them they were low level ambient environments.

At the time I had just spent a year in one location documenting one sound habitat, the Goldfields box ironbark at my acoustic sanctuary making an online twelve month sonic calendar. Spending time to know country over the long term is a requirement in an ancient sung country such as this one for me personally.

See www.hearingjaarajaara2013.wordpress.com I had been dealing with the disembodiment of sound on line in this project already. Technology would have to create a fictional and sonorous auditory landscape.

There was also the problem of not being there, on site to understand the relationship of the sound to the habitat that generated it. Nor would I be able to hear the premiere of the work or have knowledge of the acoustic space it was going to be heard in despite being given speaker specs, 4 Meyer UPJ1 loudspeakers, 1 sub to be augmented to 6 with 2 Genelec 8020s backup. The piece was to be 15-30 minutes. There could be live elements. Live Tarhu bowed spike fiddle would have been attractive to me if I could have attended. It would be an electro-acoustic six channel piece. Leah Barclay with whom I have collaborated over many years, offered to be my surrogate overseer as she would be present and was aware of the needs of my work from the outset. This was useful as in the end the files I chose from the data base she had recorded. We met and discussed this process intimately. I was feeling more connected to the site and the life of the birds. My particular focus on the sound quality and spatial movement in the acoustic space is paramount. Both were problematic in this invitation.

Once I committed to the golden eagle travelling from one habitat to another through their ambisonic recordings as the main focus of the piece, *Raptor* was born. The movement of large birds, the dihedral angles of their flight and the subtlety and strength they bring to every dramatic journey they make, the design became easy. The sound source, the calls of the golden eagle, the *Acquilla Chrysaetos* would hold the sonic identity of the piece as they move the distance between speakers in the overall migration from the Joshua Tree Creek Biosphere where it was recorded to the Beaver Creek precinct, a different habitat. To this end the speaker array was clearly conceived from the outset. In a clockwise position the design needed to take into account the northern hemispherical orientation.

SOUTH

1 **2**

EAST 6 **3 WEST**

5 **4**

NORTH

Figure 1. *Raptor* Speaker Array Orientation

The programme note describes the intention.
"The granulated eagle calls are stretched to represent the psycho-perceptual orientation of the eagle, solitary, looking down over the land. The harmonic drones of the bowed tarhu within this soundfield position the eagle's eye creating a moving point through which we can share the dynamic movements through these spaces as it rides and tips through the dihedral angles intrinsic to its flight. The slow strength and power and control of these magnificent birds can defy our aerodynamic understanding as they sculpt the biosphere.
In order to help me effect this difficult task I engaged my engineer Jim Atkins to record me bowing the tarhu in this winged like way being a bird flying and to decode the different ambisonic and stereo files I had chosen so we could effect the spatial migration. I added air sounds of the slide whistle which I recorded and granulated and for the golden element chose some ancient Greek bronze beaten cymbals to give an ancient golden feeling to the bird's DNA and latin name (Chrysaetos). We worked together on the EQ, Signal processing and effects around the east west axis of this 6 channel design. The final mix was done in the entire set up where I could tweak and add things straight into the mix live until the end. This responsive way of working with the array is not something everyone gets all the time especially with such a great engineer on board. For me I was flying on the tarhu, bringing the bird to life in the space, a long journey from the American recordings. In this situation I felt all the work was worth it and the sonic results thrilling. The sound as mass was flying around me and I was on a long journey with lots of air support.

As it happened the entire concert venue in America was changed the day before the performance, as it was felt the array of speakers was too narrow, by those at the festival. The venue was the Arizona State University Art Museum. The date, October 17, 2014. I was in Chiang Mai for a wedding and freaked out that the scale and design elements I had worked so hard for were to be altered. I emailed Jim whose speed and skill in adapting the piece into the new format was exceptional. The performance and all spatial sound considerations had to be re-jigged in 24 hours. This was of great concern to me as the entire spatial design I had spent weeks doing had to be rescaled. Jim and I had to reverse all the files reconsider the spatialisation and reverse the central trajectory had altered. It made me think, a piece is never finished. This change hit at the main design element of the piece. The axis of the flightpath. But I knew my fall back position at the moment of replay would be safe in Leah's hands. It was the first item in the concert of works with Douglas Quin, Leah Barclay, Ricardo dal Farra and Garth Paine. I would have loved to hear these composers' solutions to the invitation, "listening to life now" (dal Farra). The feedback was great. I have yet to hear it performed and look forward to it here at this festival. There will also be other opportunities.

Raptor 2014, an electro-acoustic 6 channel work, was recently acquired by RMIT for the international collection of soundart, along with the 8 channel *Loops* 1983/95.

3. CASE STUDY 2 : FIVE FACES OF MEDEA

Commission for solo performer, surround sound,
5 videos and live electronics.

Five Faces of Medea, (September 30, 2015) was written for and premiered by recorder professor virtuoso Ulrike Volkhardt, Chair of Mixed Media, Folkwang University, Essen who was invited to come to Australia for the Victorian Recorder Guild's 2015 spring festival in October. Ulrike wanted me to share a concert with her, *Traces* to open the festival and write her a new piece on the Medea myth. It's not my favourite myth (she commits 5 murders, 2 her own children, is full of hate

and gets away with it). I had to find a way in which the recorder could do justice to her sorcery and foreignness. Having met Ulrike at a recorder festival many years ago I felt I should understand more about why she wanted to do this and the skills, sounds, tools and processes which might be appropriate at this moment. I also wanted the work to be approachable for new players of electro-acoustic music and decided to implant the electro-acoustic soundfield in five video clips each one representing a key emotion or face of Medea which could be played separately or together.
The five psychological movements gleaned from the Richmond Lattimore text of the Euripides play gave the structure to the piece. The ancient Greek words, the videos, the notated parts and the electro-acoustic score work to the common goal for each movement. The intention is clear.

4. THE STRUCTURE OF THE SCORE

FIVE FACES OF MEDEA
ROS BANDT

MEDEA 1 CLIP NO 1 FOREIGNER 2'16

1. βάρβαρος (barbarous), non Greek, barbarian χένη (χenay),foreigner
Medea arrives on the shores of Corinth with Jason. Everyone looks at this exotic woman from afar. Begin the mode in the additive style of the alap, E D E G# E ...with wily turns of phrase and breath to give an exotic feel. At 50" Medea realises she is a foreigner and becomes more assertive and takes centre stage throughout. Watch the waves of the sea in the placement of the space notation.
Blow across the recorder finger holes in the direction indicated.

MEDEA 2 CLIP NO 2 MAGICIAN 2'20

2.γοής (goes), sorcerer, σοφή (sopay), clever schemer, skilled,
πανούργα(panourga),wicked μάγισσα(magissa), magician
Medea is a very talented woman capable of seduction, (slow dance) tricks, clever schemes and sorcery. (footbells). Her transformative powers are seen in the video as the sound in the water from the hydrophone makes the light patterns visually. Live effects and improvising on the key word modules give immediacy to her alchemical powers through which she ensnared Jason.

Footbells
Flattement, finger vibrato on one pitch like throbbing
Free improv. imitating ancient Greek word form and intonation

MEDEA 3 CLIP NO 3 PITY 2'37

3. δύστηνos ητιμασμένη (dustenos etimasmeni), unfortunate, dishonoured
Φέυ (Pheu), pity me
κάκιστος ανδρών (kakistos andron), worst of men
Low sad drones of video underpin the tears of rain and mock chant as she is abandoned for the local king's daughter. Sing and play together. Woe is me.

MEDEA 4 CLIP NO 4 REVENGE 3'09

4. πάς δόμος έρροι (pas domos erroi), let the whole house crash
The door slams and you improvise ff with the fire and effects full on throughout as the body/bodies burn. Cut exactly with video at 1'49. Silent freeze till the Childrens are stabbed. (offstage).You breathe on the other side of fate through the labium from 2.16. If string, breathe in Microphone with the bellows. Can also be acoustic.

MEDEA 5 CLIP NO 5 UNBEARABLE 2'07

5. δεινή, (daynay) unbearable monster
You have become something else, the hissing serpent, you have lost your original voice (no tone). Find cicada like frequencies with rocking hand over labium and altering breath pressure. Pathetic, soft, no remorse. Dry. Hot. Air is given to you (solar powered bellows) as the sungod catches you up and speeds you to the heavens in the air of the metro and aeolian harps. Your part-being goddess saves you.
1'20 walk out breathing over the fingerholes.
If strings, make vocal air sounds and disappear.

5. DESIGNING PERFORMER RELATIONSHIPS

Live electronics are also possible with the amplified recorder particularly in the second and fourth movements, Sorcery and Revenge. In Hannover I collected the sound palette and began the collaboration. I recorded Ulrike speaking the ancient Greek key words which were talismans and rhythmic DNA for the piece. These would be mixed later in Athens with ancient Greek and modern Greek voices making the chorus. (Arthur McDevitt, ancient Greek and Dora Papaioannou, the Athens Centre). I recorded her 8 foot renaissance instruments she would be unable to bring with her. I tried rushes of the composed mad dances with bells on her ankle and heard her improvise piteously and wildly. She was passionate that Medea’s voice should be just the Renaissance Tenor at A=466, so I immediately abandoned my intended use of the sopranino for the children’s death screams, (Movement 4) making them electro-acoustically instead and having them heard from speakers installed on the balcony high up behind the audience.

Much of the electro-acoustic sound and soundscapes for the 5 videos for each section for *Five*

Faces of Medea was recorded in the Mediterranean near Corinth where the myth is situated. Medea comes from the Colchis. Movement one was a Cretan sea recording where Jason's boat would have passed on its way from the east Black sea, taking Medea to her new home in Corinth, in the Peloponnese. The second movement is an iPhone recording I made of my hydrophone making light ripples in the red water tank with a MAX patch supplied by Jon Drummond and activated by Sophea Lerner. It is the score for an alchemist, translating sound and movement into image to return to transformed sound.

Figure 2. Hydrophone making the images in the water and Ulrike Volkhardt improvising as sorceress.

Here in the second movement she has various boxes containing the ancient Greek words and she is free to improvise on the meaning of each, stamping her foot with the bells of the notated dance and developing the modules with the effects as she wishes. It's short and compact and demandingly fast. The third movement has pitiful tears, recorded from rain in Hannover on a glass hotel verandah and the fourth, a vigorous fire burning a robe filmed close up in a fire stove in the bush.

Figure 3. Ulrike Volkhardt recorder, premiering Medea 4, Wyselaskie Auditorium, Parkville. 30/9/15

It rages during a rain storm on a tin roof. The performer has complete freedom to wreak havoc around the pitch centres in tandem with the video and moving sound and effects through the space.

The last movement uses a subtle blend of cicadas for the snake hissing, aeolian harps in the desert and the solar powered bellows of the organ powering Cage's longest piece in the world, ASAP at Halberstadt. This mix cross fades into the roar of the Athens metro taking Medea to the sun with her grandfather the sun god Helios. Each movement is notated on one A4 sheet for the recorder player. The video is the chronograph, 12'27. Performers are encouraged to learn it from memory to make the psychological pathway of Medea more dramatic and to avoid the music stand /microphone/ time-code barrier. Even Ulrike seemed to need these, as it is her common performer-mode for contemporary music performance.

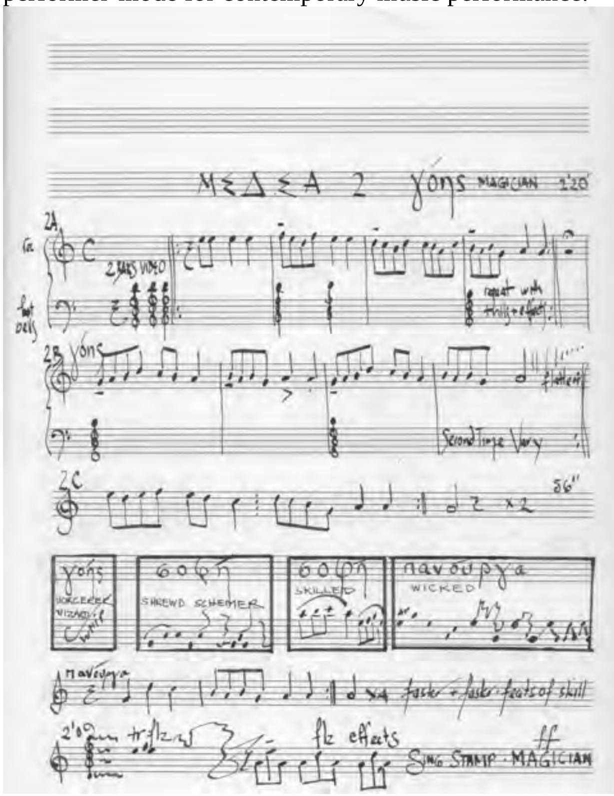



Figure 4. Movement 2: The Sorceress and Magician

Ideally the performer would use a wireless microphone to be free to move around, respond to the details of the videos, the surround sound speaker array and the audience. This yields better theatre and more integrated art, with less proscenium "performer" focus. For the premiere however, the recorder playing fraternity amongst the audience were more than thrilled seeing Ulrike being Medea, giving it her all with strong emotional content, impressive technique and perfectly timed improvisatory sections. She took the recorder to new emotional heights. The effects possible for the live performer, delay, EQ reverb, echo, distortion of the words were also prepared on the video soundtrack so that the piece could exist as a .mov file on a USB and simple notations in score. The concert performance session was automated from the Logic sessions merged with the video clips with time code for the performer to track her improvisatory pace and keep in line with the dramatic curve of the composed media throughout the five movements. An edition at A=440 is also being prepared which I will premiere with bowed tarhu recorder and voice with live electronics on the ancient greek keywords for the ABC live broadcast for the ACMA 2015 festival concert. Medea's foreignness will be further explored.

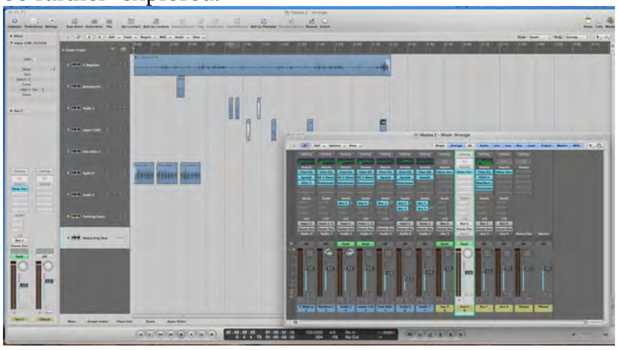


Figure 5. Movement 2 :screen shot of Logic session.

5. WHY HAVE PERFORMANCE?

The role of the performer in electro-acoustic settings can be a plus and a minus depending on the audience and the venue. I can perform *Raptor* with the array on tarhu, slide whistle and cymbals but a seat on a moving floor would also be great. If it is performed for radio, then the movement can be mixed on a slow graded right to left, south-north orientation over 15 minutes, given that it goes to air on stereo FM. This needs adjusting in the northern hemisphere.

If RMIT intends to re-situate the work as an installation in their new an outdoor sloping ground art in their new installation facility, The Bundoora Spine Soundscape System, the relationship of the height of the speakers, sound levels, EQ and distances will need to be re-set against the prevailing soundscape of the site.

Figure 6. The RMIT Bundoora Spine Soundscape

If it is a continuous loop, then boundaries need to be preserved or the sense of journey would be distorted. If it is in a stream of pieces by other people there should be a major gap. It is hoped that any further curatorial settings of the piece will be done in consultation with the composer. Raptor will always need a re-tweak in every setting for the piece to have its intended effect, especially for the geographical east-west axis to be maintained.

6. COMPARATIVE OBSERVATIONS

In these two works the constraints accepted by the commission in terms of venue, possible sounds, playback conditions with the audience have been noted.

For each piece elements need to be renegotiated with each recurrent performance, the position of speakers, the levels, the position of the audience in relation to a video screen, an outdoor concourse which is concave, fixed or moving listeners, the type and skill of the performer. Would a male Medea be possible as we have seen with a female Creon in Antigone recently?

In these two pieces, the sound transmits the emotion through many different computer processes at various stages of the composition, editing the soundscape recordings, equalising, and filtering many of the samples recorded and adding effects to them. Granulation and stretching allows the sounds to be opened up be lot, be it a bowed string of the tarhu or the call of the golden eagle. Thanks to Barry Truax, we can climb around inside the sounds in great detail.

The multi-tracking and temporal adjustment of both pieces of course lies in the computer in both cases. The array of the speakers and the export feeds to each are the critical design elements in both pieces designed at the outset, as are the sonic gestures contributed by the performer both before the mix and during performance. Finally, the spatialisation, type of speakers and their position in the listening environment in relation to the audience is the most important feature in rendering the work to the ear. Levels must be tried to enact the most desirable sculptural sounding so that the chunks of sound as mass and tiny detail occur in the right place.

7. WHY HAVE AN ENGINEER?

Since I have been performing in many of these pieces and installations over the past years, (*Floating Glass*, 6 concerts, *Iso Nagecki*, 2015, *Medusa Dreaming* 2015, I have been working in collaboration with sound engineer Jim Atkins to ensure the sound and embodied architectures are rendered during the performance with a vibrance and volatility embracing change and realtime mixes I can enjoy on the fly. He is also part of the compositional process in some of the works, so he is totally involved in the intention and construction of all elements of the work by the time we have prepared the session for the performance. Having worked with him in the ABC doing large scale inter-continental live feeds (A Global Garden for Percy, Grainger Museum), no challenge is too great. Jim has also recorded and mastered the new baroque/modern *Birdsong* CD, touring Castlemaine, Italy, Malta and next week in Elder Hall Adelaide. In the Australian concerts, I have chosen to situate the amplified historical instruments in sound fields of Australian bird recordings with two new pieces *Peristeria* (doves) and *Lyrebirds* for the Elder Hall, November concert, 2015. The engineer is an essential part of the team for this to occur. It is often challenging for the baroque musicians who prefer their sound acoustic rather than electronically mediated, but they need to be amplified to live in the surrounding digital bird speaker forest particularly when the performing venues are dry such as the historic Castlemaine Market.

Over the past forty years I have constructed some 45 installations on four continents and published many CDs of audio works. The tangos in these two works are by no means a summary of my approach in my long career. Many studio situations, indoor and outdoor, public and private, large and small scale have been used on and off site. Seven unesco world heritage sites have been installed. Ancient archeological sites bring other restrictions to the possible uses of technology as do remote off line environments such as at my acoustic sanctuary. Solar powered generators and mobile units created for the situation remain the most subtle and environmentally sustainable.
(Available at URL www.hearingjaarajaara2013.wordpress.com)

Figure 7. Flexible systems and mobile studios**.**
Working with Jim Atkins mobile system on *Windharps Symphony* in my tiny bunker for the International Harp conference 2014.

Often I have done nearly every stage of this myself, recorded sound, notated it, invented the sculptures, the multi channel playback system, rigged it, installed it, engineered it, performed in it, recorded it. Filmed it. This has given me a polymath experience as a sound artist and designer and technician. The design process is the issue. I have been fortunate to work with the best engineers in the World through commissions at the WDR Koln, the ORF Vienna and the ABC, Sydney and have worked in collaborative teams which have built works which can transcend the sum of their parts. The two works discussed here have required their specific orientation because of the nature of the details of the invitation, which can never be underestimated. I personally find a tango with an empathetic brilliant engineer frees me to participate at the moment of audition, sound painting in realtime in the electronic auditory environment I create. Both these works testify to this experience working with Jim Atkins.

It is great to see more exploratory listening spaces becoming more frequent in the computer music scene and seeing performers such as Genevieve Lacey, now designing collaborative installations combining new electro-acoustic commissions with her Dutch Van Eyck recordings in garden settings, 2015 and 2016. The context of the listening environment has a profound influence on what we make. Every space is an acoustic space.

8. MAKING WORK: WHAT IS THE FINAL AUDIENCE EXPERIENCING?

The audible space and the audience experience must reign supreme for the design of every work to be fully realised, be it an electronic concert with or without performance, or interactivity for the audience or the performer. In the two works discussed here audience interactivity is not part of the design.Temporal designs of different lengths and various forms of relationships, such as walking through a beautiful garden, may shape where the audience may want to be, and for how long they may care to listen. Babbit's old adage is still relevant, who cares if you listen? And the soundscape is playing out in its manner of operation every second. (Cage). What do we offer the auditor when making work? The question mark is always there. WHY? FOR WHOM? WHEN? HOW? and FOR HOW LONG?

Figure 8. Peters Edition plaque of Cage's quote.

9. CONCLUSIONS: CREATING WORK: THE WHY FACTOR

There are many ways to create revelatory sonic experiences for people to experience with all their senses, as these examples show. Installations, concerts,
Radiophonic works, outdoor environmental events and multi-site online works all have their power and uses. Techniques and computer possibilities are constantly changing and seducing us. Technology has always been seductive at every stage of the musical process as I pointed out in an early article in Meanjin, The musician and the Machine. (Bandt,1988). For me personally, the first shotgun microphones, the VCS3, Nagra taperecorders and the fairlight were followed by the 3DIS Interactive space, the SSIIPP 8 channel interactive

system, (Bandt 1985), sonic solutions, Audio Mulch and an array of interactive audiovisual softwares which often seem like fetish objects. But it is the quality of the intention, the design of the work and its realisation that matters. Who is it for? Where? When? For how long? Computers are helpful tools through which these questions for each work can be realized, but the presets of the softwares should never and will not dominate or create great sonic art. It's the service to which they are put through the design of the why? factor and the technique through which it is rendered that will make a work inspirational, refined or not. It has to be more than technical display. The listener is the final arbiter when the sound is physically heard and emotionally experienced in the auditory environment however diverse that may be. (Bandt, Duffy MacKinnon: 2009)

10. REFERENCES

Babbit, M. 1958. "Who Cares if you listen." *High Fidelity*. Available online at URL isites.harvard.edu/fs/docs/icb.topic244629.files/W ho%20Cares%20if%20You%20Listen.pdf

Bandt, R. 1988. "The Musician and the Machine" *Meanjin,*47(2): 274-280.Available online at URL search.informit.com.au/documentSummary;dn=578429994572457;res=IELLCC ISSN: 0025-6293. [cited 15 Oct 15].

Bandt, R., Duffy, M. and MacKinnon, D. 2007/9. *Hearing Places*: *Sound Place Time and Culture*. UK: Cambridge Scholars Publishing.

Bandt, R. 2013. "Artifice et integration dans la creation des installations sonore et musique environementale in situ: Floating Glass et JaaraJaara Digital Acoustic Sanctuary"
*Proceedings of the 2013 Colloque International: Music and ecologies of sound: Theoretical and practical projects for a listening of the world***.**
Paris: Universitaire 8. Available at URL
http://www-artweb.univ-paris8.fr/spip.php?article1677

Cage, J. 2012. "An Autobiographical Statement ", *Centennial Archives.* Available at URL johncage.org/autobiographical_statement.html

Truax, B. 1990. "Composing with real-time granular sound." *Perspectives of New Music* 28(2), 120-134

UNFINISHED BUSINESS: SOME REFLECTIONS ON ADOPTING A PRACTICE-BASED APPROACH TO TECHNOLOGICAL RESEARCH AS AN ARTIST

Rodney Berry, Ernest Edmonds, Andrew Johnston
University of Technology Sydney
Creativity and Cognition Studios

ABSTRACT

This paper reflects upon aspects of my experience as an artist moving into research and my attempts to reconcile the two areas of activity and interest. It describes several tabletop augmented reality music systems in the context of my experience as an artist working in technology research environments. It is my intention to show how the relationship between creative practice and technological research has changed for me over time and how I have come to embrace a practice-based approach to research where creative practice takes a central and crucial role.

1. INTRODUCTION AND BACKGROUND

As first author, I write in the first person from my own experience and practice. However, the co-authors acknowledged above have provided discussion and mentorship that affect the form and content of this writing.

In 1998, I had an exhibition in Paris as part of the Virtual Worlds conference. The Director of the Media Integration and Communications Lab at Advanced Telecommunications Research (ATR) in Kyoto visited the show and talked with me at length about the work. He later invited me to come to ATR as a visiting researcher. It was there that I became involved with augmented reality interfaces that had both tangible and visual elements. The first such project was *The Augmented Groove* (Poupyrev et al. 2001), a demo developed in collaboration with Ivan Poupyrev and others. We presented it at Siggraph 2000 Emerging Technologies exhibition.

1.1. The Augmented Groove

This work consisted of a stack of old vinyl phonograph records, each printed with markers that could be identified and tracked by a computer. Each record, when visible to an overhead camera, caused a particular techno music track to be played. The position and rotations of the record controlled various live effect parameters such as filter sweeps, distortion, echo, reverb etc. When more than one record was visible to the system, the looping tracks could be mixed and effected live by several people at once.

The Augmented Groove explored representation and control where the player held something tangible and each axis of translation or rotation clearly controlled a separate effect parameter. For example, rotating the record clockwise and anti-clockwise changed the sound level of its particular track whereas lifting it high above the table would add distortion or other comparatively extreme effects. The former parameter change is persistent because a record can be left on the tabletop in its rotated state. However, the latter vertical movement requires physical effort to sustain by lifting of the record, so its corresponding parameter must eventually return to its state when left on the tabletop. It made sense to assign a more extreme effect parameter, such as distortion, that was interesting in moderation but not sustainable over time. In this way, the logic of the body and the physical space, including the tabletop, became part of the program by providing distinct affordances and constraints that mirrored their role in the actual music being made.

Figure 1 The Augmented Groove (2000)

As Paul Dourish points out (Dourish 2001), a tangible interface is rarely purely tangible. The very fact that tangible elements are physically present in the real world makes them also available to our other senses, especially the visual. In addition, fiducial tracking, via the Augmented Reality Toolkit (Kato and Billinghurst 1999), added extra visual elements to the interface. The view from the overhead camera was displayed on a screen at the rear of the table-top interacting space and animated 3d CG objects and characters appeared to be attached to the records via their visual markers. This a standard way of

 ISSN 1448-7780

working with camera-based augmented reality, in this case the Augmented Reality Toolkit, to create a mixed field of view containing both real and virtual objects.
The same augmented reality technique, along with overhead camera, table-top interactions and large screen placed behind the table, was used in The Music Table project (Berry et al. 2002).

Figure 2. The Music Table (2002)

1.2. The Music Table

The user or player of The Music Table creates melodic patterns by arranging blocks on the surface of a table. The overhead camera identifies and tracks the marked blocks and uses their positions to control a MIDI synthesiser. The position of blocks on the vertical axis (toward or away from the player) determines the pitch of the note associated with that block. Its horizontal position (to the left or right of the player) determines its position on a repeating timeline. In effect, the table-top and blocks act as a musical score in the physical world.

Visual augmentations, visible on the screen with each block, add information about the length and loudness of each note. Note lengths are changed by momentarily tilting a block to the left or right in order to increment or decrement the duration of the note. Each note can be made louder or softer by rotating clockwise or counter-clockwise as one might a traditional audio volume control. Because the timeline is relatively short, the system is suited to styles of music that involve repetition, such as minimalist instrumental music and dance music.

The project was shown at SIGGRAPH 2001 Emerging Technologies Exhibition as well as a number of other similar exhibitions attached to technology research and development conferences.

Both Augmented Groove and The Music Table were designed in such a way that they could be learned quickly and enjoyed by a naïve user. This reflects the context of both as primarily technology demonstrations for the SIGGRAPH Emerging Technologies exhibition, where the expected period of engagement is between thirty seconds and three minutes before the visitor moves on to the next thing in the show.

This also reflects the roles often assigned to artists and designers in a technology research environment.

1.3. Roles of Artists

A fortunate artist might be engaged as an ‘artist in residence’ and allotted budget and an army of underlings to do their bidding. This happened to some extent at ATR and other places where Japan’s bubble economy and the technology boom in Silicon Valley contributed to a brief ‘golden age’ for media art in Japan and elsewhere. Research publications from these projects were often written by the artist’s technologist collaborators or perhaps co-written with some input from the artist. More often than not, however, artists become decorators of other people’s research. After the lab changed direction and the Art and Technology Project was no more, I and the two Japanese artists (who were employed as graphic designers), were largely tasked with creating compelling demonstrations for events like SIGGRAPH. As I became better at writing research papers and publishing, I (and other researchers) coached the other artists to do likewise. One of the artists is still working at a similar lab, but is now officially a researcher and publishes her own papers.

It is worth underlining the fact that I became a researcher through a chain of unplanned events and found myself in a community of practitioners who had already completed PhDs (or were in the process of doing so) and were socialised into the culture of research. It took some time for me to learn what was meant by such words as ‘evaluation’, or even ‘research’. It took even longer to understand that the meanings were local to that community of practice and that I needed to learn the methods and metrics recognised by this particular community enough to be able to run my own experiments, analyse the results and publish the work.

1.4. Paradigms

We were conducting quantitative research according to a particular paradigm that felt at odds with the way I naturally work creatively with technological resources. Also, I felt that researchers around me were telling stories about their research that fit a particular narrative curve. The curve starts with a clearly defined problem; a review of literature; the formation of hypotheses from the literature and the testing of these hypotheses followed by evaluation. In reality, I observed them working in a much less rational way but they later altered the story to fit the classical scientific narrative. They were not being dishonest, just inclined to forget the parts of the process that do not fit in the narrative. I suspect that a large chunk of science goes undocumented because of this.

Thomas Kuhn (Kuhn 1970, pp. 43 - 51) stresses the importance of underlying paradigms and how they serve to unify particular research communities in the sense that such communities identify with a particular paradigm. Different research communities have different levels of agreement about what is interesting and how it should be investigated. In some, there is almost universal agreement and there is little discussion of research paradigms since the paradigm is largely implicit through one’s membership of that community. For this reason, they have little need to examine the underlying paradigm or its philosophical basis. Vaishnavi and Kuechler (Vaishnavi and Kuechler 2008, p.8) describe such communities as tightly paradigmatic communities. Research into tangible interfaces and their use of representations draws on literature, theory and practice from a variety of communities that, as a research community, would best be termed "pre-paradigmatic" or "multi-paradigmatic" since, although its members might identify as a community, they are not necessarily bound together by a particular phenomenon of interest and accepted methods. This meant that, although the community in which I was located held to a particular paradigm, my field of study as a whole represents a number of communities loosely combined enough to have porous boundaries and to be open to new approaches.

1.5. Practice-based Research

As I became involved in the work of the Creativity and Cognition Studios, I began to learn about approaches to research that place creative practice firmly at the centre of the research activity. I was motivated by my previous experience to revisit aspects of my earlier work and examine it through a fresh set of cultural lenses. Eventually I decided to return to study in CCS's doctoral programme. In effect, I feel there is unfinished business in the earlier work that I believe can be resolved by adopting a different research approach.

Edmonds and Candy (Edmonds and Candy 2010) present a model (see Figure 3) for practice-based research consisting of three domains of activity: Practice, Theory and Evaluation.

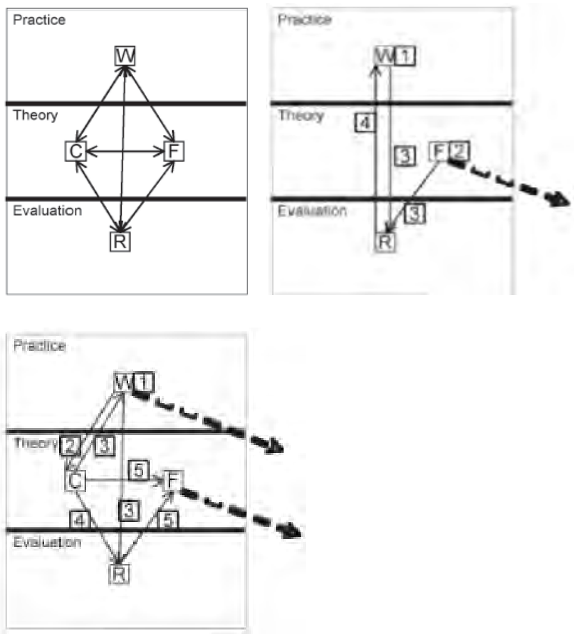


Figure 3. Three different trajectories for PBR

Practice results in Works (artefacts, installations, exhibitions, performances, etc.); Theory gives us Criteria (**C**) (design rules, guidelines etc.) and Frameworks (**F**); and Evaluation (**E**) produces Results (**R**). All are influenced by one another and the researcher might start at any point and proceed to any others, depending on their focus. In practical terms, Edmonds and Candy identify several typical trajectories based on their observations of Doctoral students at CCS. For example, one might start with theory via practice and evaluation to test and modify the theory. Or, one might start with practice and move through evaluation and theory-making in a number of iterated loops.

My chosen approach is closest to the second trajectory and my goal is to produce theory and design criteria out of an iterative process of prototyping and testing of tangible music systems with visual augmentations.

My starting point is a tangible system, very much like The Music Table, that uses the arrangement of blocks on a table-top to create *bytebeat*, a relatively recent computer music practice that traditionally relies on the computer keyboard. The focus is on capturing and reflecting upon very early stages of an iterative prototyping process whose ultimate goal is not a 'usable' commercial system. The system is developed purely to uncover design criteria and middle range theory (Merton 1968) for similar systems in the future. By middle range theory, I mean substantive theory that applies specifically to the situation at hand and is less concerned with broad applicability. Put another way, I am prepared to sacrifice a certain amount of traditional scientific validity in order to gain specific relevance to the subject of my study.

At this stage of my process, I will iteratively prototype, reflect on what is done and found, and analyse these reflections in much the same way as I will later in the user studies with more mature versions of the prototypes.

The product of the research will be presented as an art work in a variety of contexts. One context will be that of an installed work in an art gallery. Another form of the work is as a live performance instrument where the user forms and modifies bytebeat strings on-the-fly.

2. BYTEBEAT

Bytebeat is a genre of electronic music, made by forming mathematical expressions where the symbol, ***t*** represents an endless looping count of integers from **0-255**. The solution for each new value of ***t*** is fed to the computer's audio output via whatever housekeeping code is required by the programming language and its host system. For example: ***t*** on its own will produce a rising sawtooth wave, whereas ***t * 2*** will produce the same wave an octave higher (because the frequency is doubled). Traditionally, it is integer-based and uses a C-like infix syntax for the expressions. However, many also use a postfix or Reverse Polish Notation syntax and there are variations that use floating point arithmetic instead (*floatbeat*).

Given the immense space of possibilities, the resulting sounds are often surprisingly musical, if a little nasty. I speculate that this is because integers are more likely to yield frequency relationships in simple whole-number ratios, corresponding to the natural harmonic series. Simple ratios will also make rhythmic events more danceable – although the number of legs the dancer needs may vary.

Bytebeat first appeared in 2011 with a YouTube video, subsequent blog posts and short paper by Ville-Matias Heikkilä (Heikkilä 2011), otherwise known as Viznut.

The idea quickly spread and there was a flurry of activity to discover and share new expressions. Programs appeared for just about every platform, including browser-based ones that made bytebeat even more accessible.

Aesthetically, bytebeat practice tends to go in two directions. The initial push at the time of its discovery was, and continues to be, the challenge of making the most interesting musical result from the shortest string of symbols. This aesthetic is shared by the, by then well-established, demo scene.

The second aesthetic direction for bytebeat quickly emerged as programmers began to make bytebeat programs evaluate the strings in real time without waiting for compilation of the code. Changes made on the command line had immediate musical (or at least sonic) consequences. If the first direction is concerned with composition, the second direction is more about performance and improvisation. This led to bytebeat's absorption of many aspects of the live coding movement. Proficiency (even virtuosity) at the computer keyboard is part of the live coding movement's aesthetic and certainly applies to live bytebeat performance. Copy, Cut and Paste, as well as repeated Undo operations with the keyboard, allow for quick changes and recall of earlier states.

The live coders' issue of what to show on the big screen drew practitioners' attention to how the interface looked to the audience. It was also quickly discovered that the process could produce visual patterns as well as sonic and has even been used to drive a Tesla coil (Montfort, Stayton, and Fedorova 2014). Because bytebeat uses 8bit sound, it made sense for the visuals to be 8 bit as well, or at least evocative of early video games and an overall glitch aesthetic prevalent at the time.

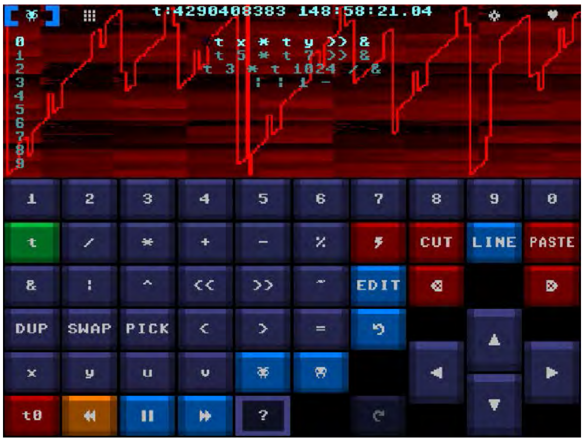


Figure 4. Screenshot of iOS bytebeat application: *The Glitch Machine* by Madgarden

Bytebeat programs, at the time of this writing, can be found for most platforms. They typically offer a keyboard interface or, for phones and tablets, a set of touch buttons for entering strings. The look and feel is often that of early video games with coarse, pixelated graphics and a limited palette of intensely saturated colours. In addition, most have some kind of representation of the waveform and/or spectrum of the output audio. The overall effect is one of sensory overload and a riot of noise and colour.

An important part of the bytebeat experience is that, although simple multiplications and divisions of ***t*** produce fairly predictable results, it is hard, even for an experienced practitioner to predict the results of an unfamiliar string. Having only existed as a genre since 2011, there may still be room for newcomers to contribute and perhaps discover something as yet overlooked. The practice is more akin to circuit-bending where one does not necessarily need to know what one is doing to discover interesting sonic possibilities.

3. TANGIBLE BYTEBEATS

Figure 5. Tangible Bytebeats screenshot

Tangible Bytebeats consists of a tabletop and a collection of 45mm cubic foam blocks, each marked with fiducial markers. An overhead camera is set up so it can see the working area and is connected to a computer. A monitor sits on the tabletop so that the markers and augmentations are visible together on the screen, along with the user's hands when manipulating the blocks. The physical set-up is similar to the author's Music Table project. The software is developed using C++ and TouchDesigner.

TouchDesigner is a procedural flow-chart-style development environment for real-time interactive 3D graphics. It has a collection of operators, represented by rectangular tiles on the screen and connected by graphical wires.

There are operators for building procedural 3D models and surfaces; lighting, texturing, rendering and compositing; manipulating live 'channels' of real-time data; creating and modifying data tables and text (including Python scripts and GLSL shaders) as well as handling data from other programs and devices. A separate program (Berry et al. 2008) uses ARToolkit to identify and track markers. The tracking data, along with the camera image and lens distortion data, are sent via shared memory to TouchDesigner where the actual modelling, rendering and compositing take place. Programming of all interaction also takes place in TouchDesigner.

The graphical display elements are currently simple cubes (Figure 5) that each correspond to one of the markers.

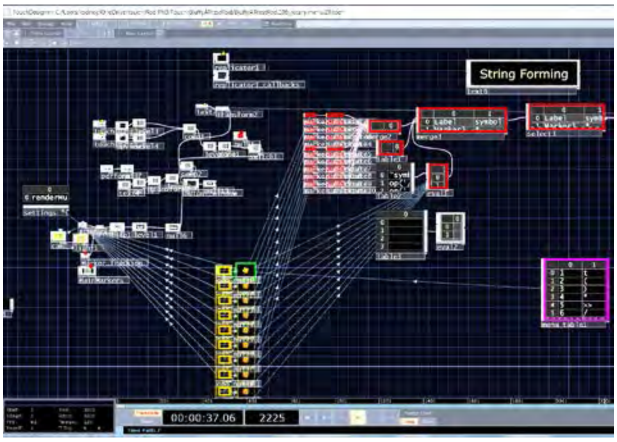


Figure 6. TouchDesigner Screenshot

Although the cubes follow the xyz translations of the markers, they do not rotate with the markers but remain square-on in relation to the camera, directly above the tabletop. Currently, each marker block has two modes of display corresponding to the two modes of function for each block. In Operator Mode, the block shows one of the following bytebeat operators:

t + - * / >> << ( ) % ^ & |

They are selected by rotating the block clockwise or anticlockwise on the tabletop.

In Number Mode, the block displays an integer between 1 and 256, once again, numbers are incremented and decremented by rotating the block in relation to the tabletop. Switching from one mode to the other is achieved by raising the block up toward the camera and back down again.

This allows the user to arrange blocks on the tabletop to form bytebeat expressions such as **t * t >> 20.** The strings are displayed at the top of the screen and also sent to a bytebeat generator, also built within TouchDesigner. A valid expression will generate sound from the computer’s audio output. At present, the system does not support the full bytebeat syntax and will require a redesign of the sound generator components. It does at least work well enough to demonstrate the concept.

It is important to point out that the interest in bytebeat began as a divergence from the design for something more like the original Music Table. The idea of using the blocks to create and edit bytebeat expressions came in a perverse intuitive flash. Aesthetically, there is a tension between a wholly keyboard-based practice that is inherently nerdy and arcane, and the addition of a very kindergarten user interface to replace the keyboard. What happens to this tension as the player explores and gets used to the system is one of the stories enacted in the process of investigation.

This is also informed and inspired by Mitch Resnick’s *Lifelong Kindergarten* work at MIT Media lab (Resnick 1998) and a belief that kindergarten-style *thinking through your fingers* has a place in the adult world. I think that this is something musicians do when using physical musical instruments and that instruments often used by composers, not just as controllers but as tangible representations.

4. NEXT STEPS

Now the prototype has been through a few iterations, I am documenting my in-process reflections and charting out the emerging design criteria from this process.

An important step will be to bring in other composers and sound artists to play with the system. This research will use qualitative data collection and analysis. Rather than a large sample of naïve subjects, I will work with a smaller number of highly expert and articulate participants. The chief data collection method will be the recorded video and audio protocols of participants' reflections while composing with the prototypes and from semi-structured interviews with the participants. This approach is largely informed by Ericsson and Simon (Ericsson and Simon 1992) as well as Schön (Schön 1983).

Analysis of data follows a Grounded Theory approach based on the Constructivist interpretation of the method described by Charmaz (Charmaz 2006) and others. Grounded Theory Methods are characterized by the following:

- The goal is to build middle-range theory grounded in the data itself rather than from the literature.
- Analysis begins as soon as the first data is collected. Subsequent collection is iteratively informed by the emerging analysis.
- The literature review proper happens after data gathering and analysis has already begun. The idea is to give the data a chance to speak and for the researcher to keep a fresh perspective on what is observed without forcing one’s analysis with predetermined theoretical frameworks. In a sense, the literature is treated more as data and less as a sacred source of wisdom.

I am using this particular approach because the study is exploratory in nature and does not readily lend itself to a classical quantitative approach. There is also a good fit between the nature of my own art practice and the prototyping process as a way of discovering/constructing and developing ideas. There is also fit between art practice and the grounded theory approach of inductive, constant comparison and movement toward higher levels of abstraction in the analysis of data. The on-going dialogue with one’s data is, for me at least, analogous to the artist’s ongoing conversation with his or her materials.

5. CONCLUSIONS

Over the years, working in creative practice and technology research has given me a kind of split personality that I am only now beginning to re-unify through a practice-based approach to research. The purpose of this writing has been to describe several similar creative projects and how my approach has changed toward the research process from which they emerged. For me, the important learning is that it is possible to do research without putting creative practice aside and that creative practice has a place at the centre of the research process.

6. REFERENCES

Berry, R., I. Poupyrev, M. Tadenuma, and N. Tetsutani. 2002. "Inside the score: music and augmented reality." *International Workshop on Entertainment Computing Technologies and Applications*.

Berry, Rodney, Janaka Prasad, Jörg Unterberg, Wei Liu, and Adrian David Cheok. 2008. "Augmented Reality for Non-Programmers." ACM SIGGRAPH, Los Angeles, August 15.

Charmaz, Kathy. 2006. *Constructing grounded theory : a practical guide through qualitative analysis*. London: SAGE.

Dourish, Paul. 2001. *Where the action is : the foundations of embodied interaction*. Cambridge, Mass.: MIT Press.

Edmonds, Ernest, and Linda Candy. 2010. "Relating theory, practice and evaluation in practitioner research." *Leonardo* 43 (5):470-476.

Ericsson, K. Anders, and Herbert Alexander Simon. 1992. *Protocol analysis : verbal reports as data (Revised edition)*. Rev. ed. Cambridge, Mass: MIT Press.

Heikkilä, Ville-Matias. 2011. "Discovering novel computer music techniques by exploring the space of short computer programs." *arXiv preprint arXiv:1112.1368*.

Kato, H., and M.; Billinghurst. 1999. "Marker tracking and HMD calibration for a video-based augmented reality conferencing system." 2nd IEEE and ACM International Workshop on Augmented Reality, 1999. (IWAR '99), San Francisco, CA.

Kuhn, Thomas S. 1970. *The structure of scientific revolutions*. 2nd ed. Chicago: : University of Chicago Press.

Merton, Robert King. 1968. *Social theory and social structure*. 1968 enl. ed. New York,: Free Press.

Montfort, Nick, Erik Stayton, and Natalia Fedorova. 2014. "The Trope Tank: A Laboratory with Material Resources for Creative Computing." *Texto Digital* 10 (2):53-74.

Poupyrev, I., R. Berry, M. Billinghurst, H. Kato, K. Nakao, L. Baldwin, and J. Kurumisawa. 2001. "Augmented Reality Interface for Electronic Music Performance." proceedings of the 9th International Conference on Human-Computer Interaction (HCI International 2001).

Resnick, M. 1998. "Technologies for Lifelong Kindergarten." *Educational Technology Research and Development* 46 (4).

Schön, Donald A. 1983. *The reflective practitioner : how professionals think in action*. New York: Basic Books.

Vaishnavi, Vijay, and William Kuechler. 2008. *Design science research methods and patterns : innovating information and communication technology*. Boca Raton: Auerbach Publications.

EXPLORING A DISEMBODIED VOICE IN SURROUND SOUND MEDIA SPACE

Dr Colin Black
School of Humanities and
Communication Arts
Western Sydney University
artist@colinblack.com.au

ABSTRACT

This paper explores ideas related to Electronic Voice Phenomena (EVP) and the disembodied voice in mediated space as a starting point to develop a 5.1 surround sound radio art work. It documents the author's creative processes and 5.1 production techniques that include a type of deconstruction of spatiality and fluctuating relationship with sonic "reality" and fiction. Moreover, the paper explores the notion of varying mediated "distance" between the listener and the recorded voice in terms of both the "length" or transformative effect of the audio chain and/or communication system and "time" that can be heard as a technologically related chronological aspect.

INTRODUCTION

> *When I turn my radio on, I hear a whole chorus of death rattles ... from voices that have been severed from the body so long that no one can remember who they belong to, or whether they belonged to anyone at all ...*
>
> - Gregory Whitehead

> *I began listening to the sounds around me from all directions and imagined how such sounds, and in just such complexity, could be transmuted into music. It excited and stimulated me to think about the possibility of such a metamorphosis and I began to imagine the invention of new devices that would make spatial music possible ...*
>
> - Edgard Varèse

This paper examines my short form work entitled, "A Disembodied Voice In [5.1 Surround Sound Radio] Space" (duration 7'16") that combines elements of Whitehead's disembodied radiophonic voice with Varèse's notion of spatial music. In many ways this work can be thought of a meta-referencing piece that explores sonic space and the potentially disconcerting schizophonic and metaphysical aspects of the medium.

BACKGROUND

On 6 December 1877, Thomas A. Edison unwittingly dislocated the voice from the body by recording his voice onto a tinfoil roll. This psychophysical complexity was further exasperated by the development of the radio, where voices seemed to travel through - or manifest from - thin air, like spirits. This perceived phenomena sparked a fascination with the idea of receiving the voices of "phantoms" on radio. Ernst Fisk in 1930, predicting "an even more wonderful [radio] age ahead of us; we cannot exhaust the infinite," (Ernst 1930a) also speculated about the possibility of talking to the dead using the wireless (Ernst 1930b). Moreover, Antonin Artaud's 1947 radio play, *To Have Done with the Judgment of God* explored the idea of "a body without organs" (Artaud 1976) and William Burroughs's work has explored the theme of the interconnection between technology and magic.[1] More recently, John Potts's radio piece, *Castlereagh Street* (2000) was based on a ghost who "haunts the corner of Castlereagh and Market streets in Sydney (Australia) and experiences the world primarily through sound" (ABC Classic FM 2000), while Allen S. Weiss has written about the "themes of modernism … occur[ing] as a matter of course in the phantasmic realm of radio" (Weiss 1995).

The advent of these new mediums (sound recordings and wireless transmission) also gave birth to Electronic Voice Phenomena (EVP), a belief that elements of the noise found on the radio [and in sound recordings] resemble "voices of the dead and other discarnate entities" (Baruss 2001).

After four years of experiments, Swedish painter and documentary filmmaker Friedrich Jürgenson detailed how he could contact the spiritual world using standard radio receivers, in his book entitled *Radio Contact with the Dead* (Sconce 2000). Susy Smith's research into the development of Jürgenson's EVP work claims that almost from the beginning of his experiments, Jürgenson had heard one insistent female voice repeatedly pleading,

[1] For example in *Cities of the Red Night*, missing persons are located via sound recordings of their environment and randomly cut in recordings of the subject's voice.

"Make contact through radio" (Smith 1977) which led him to utilise radio. Later Konstantin Raudive also used radio to listen to the dead; by tuning "from one end of the wavelength-scale to the other" Raudive claimed a voice would suddenly indicate when to stop saying "Now" or "Make recording!" (Sconce 2000). Raudive's recordings of radio spirit voices include their cries to be listened to via radio.

> Please radio!
> Bit by bit only through radio.
> Just let the radio loose!
> Speak through radio!
> Through the radio
> What a rascal, switch on the radio! (Sconce 2000)

For many, broadcasts and sound recordings seem to reach across the fabric of reality to form an interface between this world and another.

AESTHETIC AND CONCEPTUAL STANDPOINTS

While the theme of radio and the voices of the dead - or spirits - have emerged time and time again within radio works, very little radio art work (that explores this theme) has been developed to utilise 5.1 surround sound production and broadcast techniques. With "A Disembodied Voice In [5.1 Surround Sound Radio] Space" I explore the idea of a disembodied voice within the added dimension of 5.1 surround sound, as opposed to the monophonic or stereophonic "acoustic frame" or "sound stage" that has been chiefly used to date. The aim is also to create an immersive sonic environment that takes the listener inside its constructed sonic world, where the disembodied voice thematically develops the notion of the formless body and instantaneous travel via thought and desire.

1.1. Treatment

The work - while it contains sparse passages of dialogue - is not driven by text narrative, but aims to create impressions much in the same way as poetry does. It is a form of sonic poetry where language, sound effects (including location recordings) and musical composition have equal importance and counterpoint each other to convey meaning that is highly individualistic for each listener.

In this poetic style the work's use of fragmented text narrative is employed so as to help to form sections and a framework between which audio segues propel the work along new trajectories. This audio treatment and framework also assists the development of the theme of the formless body and instantaneous travel via thought and desire (i.e. the listener suddenly and seemingly emerges in the new location with the Disembodied Voice character).

The work's sound world is an aural space that is both influenced by reality and distorted by the idea of the spirit world. To reflect this imagined sonority I have blended techniques from electronic music, musique concrète and audio documentary. So while the location recordings may be encoded with natural sounds and have a coherent aural image, when these recordings are treated and/or juxtaposed in this manner, the resulting balance helped to form the work's pluralistic sound world that is the confluence between this world and an imagined other world. This was also achieved by creating an interplay between a dominance of filtered/treated location recordings, electronic music composition and the more naturalistic use of location recordings. Parts of the work also sonically superimpose audio elements in a way that would be impossible in reality.

A short audio recording that is claimed to be a genuine EVP recording published by The EVP & Transcommunication Society of Great Britain and Ireland is utilised in the work to add an "authentic" EVP element. The short sample is claimed to say "I'll go there......Hallam....the dead man [sic]" [1] and can be heard in the left and right surround speakers from 0'18" to s 1'03" in the work. Essentially to my ears this recording sounds like static that has abrupt changes in level, nevertheless this is claimed to be genuine.

Based on Jürgenson and Raudive's, work I have also made recordings of tuning a radio receiver from one end of the wavelength-scale to the other and of radio static, in the hope of making my own EVP recordings. I have used this sound as an audio motif throughout the work while also exploring Friedrich Kittler's concept that this could be essentially the purest way to listening to the radio. Kittler states:

> Nobody listens to radio. What loudspeakers or headsets provide for their users is always just radio programming, never radio itself. Only in emergencies, when broadcasts are interrupted, announcers' voices dry up or stations drift away from their proper frequencies, are there any moments at all to hear what radio listening could be about. (Kittler 1993)

As well as exploring the idea of sampling broadcast audio material and feeding this back into the work to create a type of audio media feedback or resonance when it is sequentially broadcast and further transformed by the audio chain and communication system, this work also explores contrasting low fidelity and high fidelity audio aesthetics. The intention is to highlight the fluctuating audio aesthetics of the mediated voice and to create a type of varying mediated "distance" between the listener and the actor's voice. As an artist I am intending that low fidelity aesthetics communicate an increased "distance" in both the "length" or transformative effect of the audio chain and/or communication system (i.e. how many stages or apparatuses the audio has passed through or been transmitted across) and "time" that can be heard as a technological chronological aspect (e.g.

[1] EVP recording (which can be found at http://www.ferio.fsnet.co.uk/wav/Wavweb1_2.wav) was made by Judith Chisholm. "Extract 1: Spirit Voice (male) 18 November 1999: I was alone in my room and asked my friend Jack Hallam (who died in 1986) if he would accompany me somewhere that evening and announce who he is". The EVP & Transcommunication Society of Great Britain and Ireland, "EVP Wav Extracts From Recordings," The EVP & Transcommunication Society of Great Britain and Ireland, http://www.ferio.fsnet.co.uk/WAVEXAMPLES.htm (accessed June 5, 2012).

valve or tube distortion tends towards time stamping sound recording or it sounds like an older recording).

This impressionistic production style for "A Disembodied Voice In [5.1 Surround Sound Radio] Space" with its surreal sound world, heightened sense of sonic perspective, audio segues, poetic story fragments, low fidelity and high fidelity audio aesthetics and media specific self references is explored in the work so as to support the development of the Disembodied Voice.

1.2. Recording Considerations and Methodology

A script was developed for the project and the dialogue was recorded in the Anechoic Chamber at the University of Sydney's Acoustics Laboratory (see Figure 1) to allow for maximum audio manipulation choices when later editing and mixing the recording into the final work. Some of the dialogue and sound effects were also recorded in the Reverberation Room at University of Sydney's Acoustics Laboratory (see Figure 2)[1]. The contrast between the recording styles (i.e. anechoic, reverberant and on location) helped to facilitate a sonic transformation from worldly acoustics to "other" worldly imaged sonic environments. Various surround reverb effects were also applied to the studio dialogue recordings to investigate a range of fabricated acoustic spaces in which to place the voice.

Figure 1. Yanna Black rehearsing script in the Anechoic Chamber at the University of Sydney's Acoustics Laboratory (photo by Colin Black).

[1] NB Only the dialogue recordings made in the anechoic chamber were used in the final work.

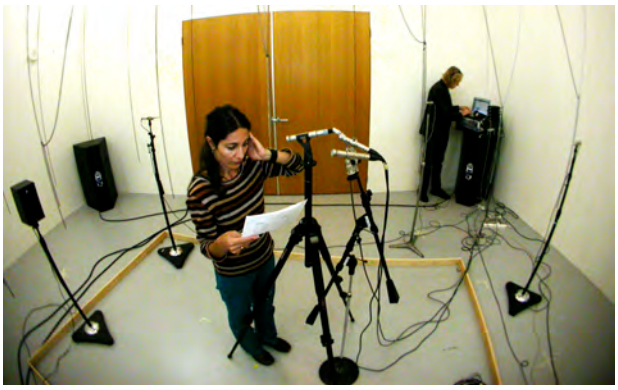


Figure 2. Yanna Black rehearsing script in the Reverberation Room at the University of Sydney's Acoustics Laboratory with Colin Black at the computer preparing to record (photo by Dr. William L. Martens).

The factors that contribute to the overall sonic aesthetic of the unfiltered components for this work include utilizing a Sennheiser MKH 40 microphone (used for the Anechoic Chamber recordings), three Oktava MK-012 microphones (for the Reverberation Room recordings), and two Sennheiser MKH 40 microphones with a Sennheiser MKH 30 microphone (for the surround sound location recordings). These were recorded with a range of devices including a Sound Devices 722 (Portable, High-Resolution Audio Recorder) and Pro Tools software with an MBox2 or 002 audio interface at a sampling rate of 48kHz with a 24 bit depth resolution and PCM encoding.

I chose to record the dialogue in the Anechoic Chamber with a Sennheiser MKH 40 mounted inside a Rycote, Windshield (WS ST AE), which is pictured above on the right in Figure 1. I chose the MKH 40 because it has an exceptionally flat frequency response from 40 to 20,000 Hz (see Figure 3) and therefore adds very little tonal colouration to the recording. In addition to this the MKH 40 has exceptionally low inherent self-noise and high sensitivity specifications,[2] which contribute towards the low level of noise/hiss in the recording of performer, Yanna Black's very quiet, whispered dialogue in sections of the work. The Rycote, Windshield was used instead of a pop shield to reduce the effect of potentially explosive "P" and "B" phonemes in the recording. Further contributing to the transparency of the dialogue recordings was the use of the Sound Devices 722 (Portable, High-Resolution Audio Recorder), which has exceptionally low inherent self-noise pre-amps and analogue to digital converters.[3]

[2] "Equivalent noise level A-weighted (DIN IEC 651)..... 12 (18) dB" and "Sensitivity (free field, no load, 1 kHz)......... 25 (8) mV/Pa ± 1dB". Sennheiser, "MKH 40 Product Sheet," Sennheiser, http://en-de.sennheiser.com/downloads/6316d267e77c057353d4ec766ac77e2e.pdf (accessed June 11, 2012).

[3] "Equivalent Input Noise: Mic: −128 dBu max (−130 dBV), 150 ohm source, 20 Hz–20 kHz BW flat filter, gain fully up". Sound Devices, LLC., 722 High Resolution Digital Audio Recorder User Guide and Technical Information firmware rev. 2.67, (Reedsburg, WI: Sound Devices, LLC., 2011), 69. http://www.sounddevices.com/download/guides/722_en.pdf (accessed June 11, 2012).

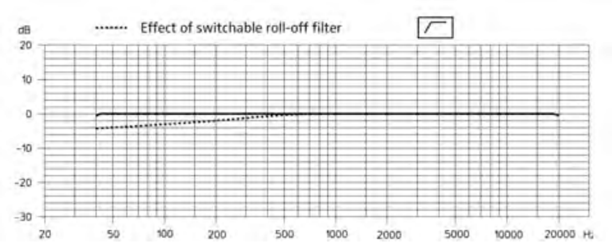


Figure 3. Sennheiser MKH 40 Frequency Response Chart. Sennheiser, “MKH 40 Product Sheet.”

I chose to record reverberant sound effects and additional dialogue in the Reverberation Room with the Oktava MK-012 microphones as these microphones have a “presence peak” (see Figure 4) and are effective in articulating the high frequencies of the early reflections in the room. Each microphone was equipped with a cardioid capsule with one microphone facing the performer, while the other two microphones were set up in an ORTF stereo configuration with the pair of cardioids set at an angle of 110˚ and roughly spaced equal to ear spacing (see Figure 2). Tomlinson Holman in his book *Surround Sound* claims that the ORTF stereo has “won over M-S, X-Y, and spaced omnis in blind comparison tests” (Holman 2008) for spatial audio representation, which was also an important consideration. This microphone set up also allowed me to choose between mono and stereo reverberant takes of the performance at various distances from the direct sound source. This set up was recorded using Pro Tools software with a 002 audio interface with its onboard pre-amps.

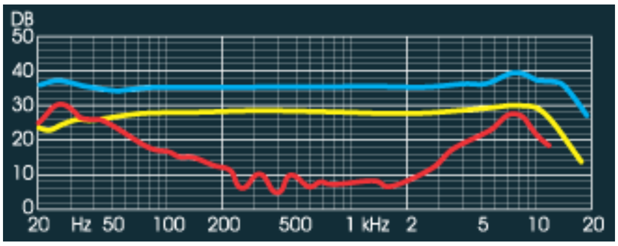


Figure 4. Oktava MK-012 Microphone (with Cardioid Capsule) Frequency Response Chart.

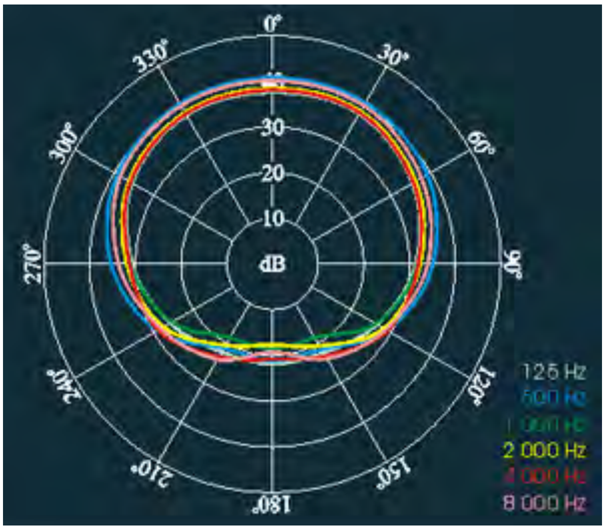


Figure 5. Oktava MK-012 Microphone (Cardioid Capsule) Polar Pattern.

Two locations (Castlereagh Street, Sydney and the interior of a Toyota Corolla) - were recorded using a Double M-S microphone configuration (using two Sennheiser MKH 40, cardioid microphones and one Sennheiser MKH 30, figure of 8 [see figure 6]) onto 3 tracks of a portable recording set up. This portable recording set up consisted of the Sound Devices 722 (Portable, High-Resolution Audio Recorder) with pre-amps and analogue to digital converters for the MKH 40 microphones, which were routed digitally via S/PDIF into channels three and four of the MBox2 and recorded using Pro Tools software running on a MacBook Pro laptop. Additionally, the MKH 30 microphone was connected to the pre-amp of channel two of the MBox2, which was simultaneously recording in Pro Tools. Notably the MKH 30 microphone has almost identical specifications to that of the MKH 40 (except for its figure of eight polar pattern) and was therefore a very suitable companion to this set up to achieve naturalistic, correlated, surround sound location recordings. This was also a practical choice of microphone set up due to the high level of portability needed to make the location recordings. The matrixing of the Double M-S set up was performed after the recording in Pro Tools, to maximize mixing options.

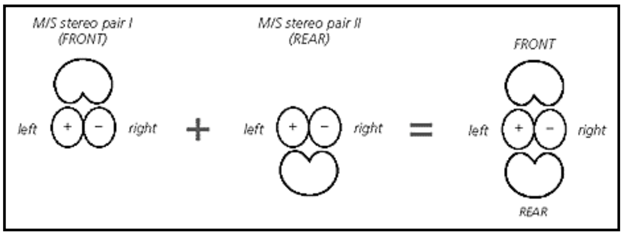


Figure 6. Double MS Recording Setup Schematic. Schoeps Mikrofone, “Double M/S,” SCHOEPS GmbH.

ANALYSIS OF THE WORK

"A Disembodied Voice In [5.1 Surround Sound Radio] Space" consists of six contrasting sections, creating surprise throughout the piece. Using the audio segues to vary the pace, together with the varying of textures, spatial aspects, the works fragmented narrative form and unusual theme make it difficult to anticipate what is coming and all work to hold the listener's interest.

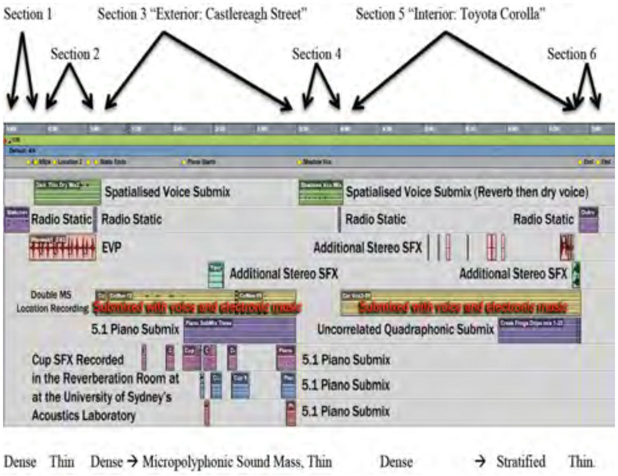


Figure 7. Screen Shot of Final Compilation Pro Tools Session (with Descriptive Labels).

The first sound that is heard in the work is that of multiple electrical switches turning on (indicative of the electricity necessary for the existence of Electronic Voice Phenomena), followed by the establishment of a dense stratified texture of spatialised radio static and recordings of tuning radio receivers from one end of the wavelength-scale to the other. Section 1 introduces the audio motif of radio static and sets up the idea of searching for the Disembodied Voice character in the radiophonic mediated landscape much in the same manner of that of Jürgenson and Raudive. It also pays homage to John Cage's *Imaginary Landscape No. 4* (1951), Stockhausen's *Kurzwellen* (1968), William Basinski's *The River* (2002), and Tod Dockstader's *Aerial* (2003), which all explore, in different ways, the sound of radio reception in their works. Section 1 then segues into our first encounter with the Disembodied Voice character in Section 2 (which has a contrasting thinner texture) using the R360˚ Surround Reverb plug-in in Pro Tools to mask the transition.

In Section 2, the Disembodied Voice is treated so as to make it sound like a low fidelity recording to highlight the electroacoustic perception that there is a type of mediated "distance" between the listener and the actor's voice, with the suggestion of being mediated via EVP from the spiritual world. To achieve this sonic quality, I used Logic Pro's Channel EQ plug-in to remove the low frequencies and high frequencies and left a narrow band around 2kHz to make the voice sound telephonic. I then passed the voice through the Guitar Amp Pro plug-in, which I set to emulate tube distortion, adding a nostalgic tube amplifier effect. I also used the Bus sends to create a wet dry mix, where the wet mix is time delayed by roughly two seconds and further processed using a Spectral Gate plug-in and Sound Designer (convolution reverb) plug-in set to a "Marble Church" impulse response. The panning of this 5.1 wet mix was then automated so as to send this wet mix to a different speaker for each phrase spoken by the performer. The aim was to create a sonically deteriorated reflection of the Disembodied Voice that "appears" to move around the space.

Also heard in the background (in the left and right surround speakers) is the looped EVP recording "I'll go there......Hallam....the dead man [sic]"[1]. Radio static is then used again at the end of this section to propel the work in a new direction.

Emerging from the abstract electronic aesthetic of Section 2 we arrive suddenly in Section 3's naturalistic, correlated surround sound urban environment of Castlereagh Street in Sydney with its contrasting thicker texture. I chose this location as John Potts's *Castlereagh Street* (2000) was based on the idea that this location was haunted (ABC Classic FM 2000) and in a way I am continuing to sonically search the site for spirits. The electronically generated sounds that start at around 1'26" into the work, during this section, were created from recordings of the Manly Ferry, travelling through the waves near the headlands in 2005, that have been filtered many times using the Spectral Gate plug-in and other audio filters in Logic Pro. I chose to filter this sound, as a number of cultures believe that the dead journey across the waters to reach the Afterlife that is either on an island or in a place under the water.[2]

It is in this section that the Disembodied Voice character starts to realize that something is wrong: "My coffee, I can't taste my morning coffee … [and] The sun's turned cold. I can't feel the warmth of the sunlight on my skin." To highlight this miscorrelation with reality, and to create a kind of heterophonic texture, the phrase "the sun's turned cold" is delayed at various times around the five speakers of the 5.1 surround system (see Figure 8).

---

[1] The EVP recording entitled "Extract 1: Spirit Voice (male)" by Judith Chisholm can be found at http://www.ferio.fsnet.co.uk/wav/Wavweb1_2.wav. Chisholm states, "Extract 1: Spirit Voice (male) 18 November 1999: I was alone in my room and asked my friend Jack Hallam (who died in 1986) if he would accompany me somewhere that evening and announce who he is". The EVP & Transcommunication Society of Great Britain and Ireland, "EVP Wav Extracts From Recordings," The EVP & Transcommunication Society of Great Britain and Ireland, http://www.ferio.fsnet.co.uk/WAVEXAMPLES.htm (accessed June 5, 2012).

[2] "The Egyptians, whose land was the gift of the river Nile, … believed that the dead crossed over water, and fashioned the hearse in the form of a boat." The 1911 Classic Encyclopedia, "Funeral Rites," LoveToKnow Corp. Inc., http://www.1911encyclopedia.org/Funeral_Rites (accessed June 12, 2012). "Irish traditions tend to envision the Afterlife as being entered through a location which is almost always off to the west (where the sun sets) and usually either an island at sea or a place actually under the water." The Druid Network, "Lesson Twelve ~ Funeral Rites and the Afterlife" The Druid Network, http://druidnetwork.org/learning/courses/online/polytheist/twelve (accessed June 12, 2012).

Figure 8. Pro Tools Session Screen Shot of "Castlereagh Street" Dialogue Surround Sound Delay.

In contrast to the electroacoustic aesthetic qualities of the character's voice in the previous section, minimal processing has been applied to the actor's whispered lines, recorded at very close proximity to lessen the mediated "distance" between the listener and the actor's voice.

The sound effects of the coffee cup, which were recorded in the Reverberation Room, are then superimposed in a spatially incoherent fashion to further illustrate the Disembodied Voice's miscorrelation with reality. We also hear the Disembodied Voice attempting to swallow in frustration.

At this point, five atonal improvised piano parts are spatially placed around the 5.1 surround "sound stage". This already dense texture of pianos is further augmented with an increasing number of improvised atonal piano parts, which are all constrained to the high register of the piano, gradually creating a micropolyphonic sound mass. In earlier experimentation with the 5.1 surround sound system I realised that this type of speaker configuration presented micropolyphonic sound masses in a much more detailed fashion than a stereo system could ever attain and this was one of the reasons why I chose to compose in this style for this project. I was also inspired by Edgard Varèse's notion of spatial sound masses, when he stated, "It was Helmholtz who first started me [Varèse] thinking of music as masses of sound evolving in space, rather than, as I [Varèse] was being taught, notes in some prescribed order" (Varèse 1968).

At the end of this section, the increasing density of the texture is suddenly released into the contrasting thinness of the texture of Section 4, again using the R360˚ Surround Reverb plug-in in Pro Tools to mask the transition.

Spatiality is further deconstructed with the reversing of audio gestures in Section 4, by having the abstracted uncorrelated 5.1 surround sound reverb (with its homogenous heterophonic texture) occurring before the monophonic originating anechoic recording of the actor's voice (see Figure 9). Dr. William L. Martens contributed to this work by preparing the surround sound uncorrelated autoconvolution reverb parts using MATLAB software, which I further enhanced, using the R360˚ Surround Reverb plug-in in Pro Tools.

The script for Section 4, that includes, "Light is like architecture; with it we define space. Without shadows light has no form." thematically questions elements of spatiality. The anechoic recordings of the performer's voice, having no audio reflections, give us no sonic clues to the size of the room in which the voice is recorded. This acoustically communicates a Disembodied Voice in a potentially infinite sized, air filled space. This section closes with the line "Are we shadows in the light? Shadow forms?" after which the audio motif of radio static returns and is again used to propel the work in a new direction.

Figure 9. Pro Tools Session Screen Shot of Surround Sound Reverb Followed by Anechoic Recording.

In Section 5, we seemingly materialise in the interior of a Toyota Corolla, with the sound of rain on its windshield, recorded in surround sound. From this interior perspective we hear the car doors open and close and someone (whom because of the context of the piece we assume is the Disembodied Voice character) fasten their seat belt, start the car and drive off. Combined with this naturalistic soundscape we also hear a recapitulation of the filtered recordings of the Manly Ferry, which were present in Sections 3. As the Disembodied Voice character begins to speak, the sound of the interior of the Toyota Corolla and the actor's voice are filtered by the GRM Comb Filter - "a filter that resonates a selected frequency and all of the harmonics of that frequency" (Chadabe and Favreau n.d.) - and the Reson filter - that "create[s] new sounds from current sounds by rebalancing and redistributing the sound's resonant frequencies" (Ibid).

At around 5'24" an audio gesture of heavy rain drops in the front speakers is filtered and delayed before the same gesture is heard in the left surround and then the right surround speaker. The intention here is to create an artificial sense of forward motion. Other similar gestural responses also occur later in this section.

Water drips that were recorded at very close proximity in an anechoic chamber are placed in the 5.1 surround "sound stage" with added cave-like reverb. I further augmented the mix with two stereo recordings of frogs and creek sounds that I set up as an uncorrelated quadraphonic submix by assigning one stereo recording to the front left and right speakers and the other to the left and right surround speakers. I then slowly increased the overall level of this uncorrelated quadraphonic submix so as to superimpose it onto the interior car "sound stage", with the intention of creating a stratified texture that has an incredulous relationship with reality. That is, the car continues to sound like it is in motion while it is progressively being submerged in a swamp like environment. Again, this links back to the belief that a number of cultures have, that the dead journey across and sometimes under, the waters to reach the Afterlife.

Section 5 ends with the actor whispering the words "The image of my body disintegrates, with my memories in the rain" and the sound of an ambulance siren followed finally by water bubbles. The idea here is to evoke a sense of emergency, alluding to the Disembodied Voice character being killed in a car accident and the final departure of the Disembodied Voice's spirit from this world.

The work concludes with the 5.1 surround "sound stage" in Section 6 suddenly collapsing with only the front centre speaker emitting the sound of radio static in a final EVP search for the Disembodied Voice. The final sound we hear in the work is that of a single electrical switch being turned off and therefore the EVP search being abandoned.

CONCLUSION

This paper augments the discourse regarding the creation of 5.1 surround sound radio art works and documents a range of the author's creative processes, conceptual standpoints and production techniques in making the aforementioned work. Moreover, by exploring ideas related to the deconstruction of spatiality and fluctuating relationships with sonic "reality" and fiction alongside an increased awareness of electroacoustic perception (that gives us sonic clues to mediated "distance" between the listener and the actor's voice), this paper contributes to the discourse pertaining to media space and the disembodied voice.

In contrast to conventional radio drama (that is typically driven by spoken word dialogue), this work's impressionistic treatment of the phantasmic narrative (where all available audio elements are given more equitable weighting) creates an inclusive delicate sonic framework that invites the listener to immerse his/herself in the works' surround sound environment. Through this immersion the line between sonic "reality" and fiction is increasingly blurred until it becomes believable within the mediated space, that the Disembodied Voice character navigates its identity via the formless body and negotiates instantaneous travel via thought and desire.

REFERENCES

ABC Classic FM. 2000. "The Listening Room (Archive Website): April 2000." Australian Broadcasting Corporation. Available online at abc.net.au/classic/lroom/stories/s113169.htm. Accessed April 2015.

Artaud, A. 1976. "To Have Done with the Judgment of God," in *Antonin Artaud Selected Writings*, ed. Susan Sontag. Berkeley: University of California Press.

Baruss, I. 2001. "Failure to Replicate Electronic Voice Phenomenon." *Journal of Scientific Exploration* 15(3): 355 – 367.

Chadabe J. and E. Favreau. n.d. *GRM TOOLS CLASSIC VST User's Guide.* Paris: Ina. Available online at dtic.upf.edu/~lfabig/resources/materials/session8-spectral_processing2/GRMToolsClassicVSTXPDoc16Eng.pd. Accessed June 2012.

The Druid Network. 2003. "Lesson Twelve ~ Funeral Rites and the Afterlife." In *Polytheist Druidry*. The Druid Network, Available online at druidnetwork.org/what-is-druidry/learning-resources/polytheist/lesson-twelve/. Accessed April 2015.

Fisk, E. 1930a. "Radio Era: Next Step – Phone N. Zealand. Fisk Tells," *Evening News*, May 1.

Fisk, E. 1930b. "Talking to Dead by Radio: Mr Fisk's Prediction", *Herald (Melbourne)*, May 26.

Holman, T. 2008. *Surround Sound : Up and Running.* Amsterdam: Elsevier/Focal Press, 2008.

Kittler, F. 1993. "Die letzte Radiosendung," in *On the Air: Kunst im öffentlichen Datenraum*, ed. Heidi Grundmann. Innsbruck: Transit.

Schoeps Mikrofone. n.d. "Double M/S." SCHOEPS GmbH. Available online schoeps.de/E-2004/double-ms.html. Accessed April 2009.

Sconce, J. 2000. *Haunted Media : Electronic Presence from Telegraphy to Television*. Durham: Duke University Press.

Sennheiser. n.d. "MKH 40 Product Sheet." Sennheiser. Available online at en-de.sennheiser.com/downloads/6316d267e77c057353d4ec766ac77e2e.pdf. Accessed April 2015.

Smith, S. 1977. *Voice of the Dead.* New York: Signet.

Sound Devices, LLC. 2011. "722 High Resolution Digital Audio Recorder User Guide and Technical Information firmware rev. 2.67." Reedsburg, WI: Sound Devices, LLC. Available online at sounddevices.com/download/guides/722_en.pdf. Accessed April 2015.

Oktava USA. n.d. "MK-012 Single Microphone." Oktava USA. Available online at oktavausa.com/ProductsPages/MK-012.html. Accessed April 2015.

Varèse, E and A. Alcopley. 1968. "Edgard Varèse on Music and Art: A Conversation between Varèse and Alcopley" *Leonardo* 1(2): 187-195.

Vella R. and A. Arthurs. 2000. *Musical Environments : A Manual for Listening, Composing and Improvising*. Sydney: Currency Press.

Weiss, A. 1995. *Phantasmic Radio*. Durham: Duke Universtiy Press.

Whitehead, G. 1991. "Radio Art Le Mômo: Gas Leaks, Shock Needles and Death Rattles." *Public,* 4/5: 140-149.

The 1911 Classic Encyclopedia. n.d. "Funeral Rites." LoveToKnow Corp. Inc. Available online at 1911encyclopedia.org/Funeral_Rites. Accessed June 2012.

ESCAPEE GLOSS: A SYMPHONY OF POLYMEDIA

Brigid Burke
Independent
Composer/Performer/Artist
www.brigid.com.au

ABSTRACT

Escapee Gloss: A Symphony of Polymedia explores the methods and meaning of polymedia processes through the composition and subsequent performance of an interactive seven-movement piece, *Escapee Gloss* composed from 2010-14. The creative work is a documentation of musical and graphic scores and polymedia recordings, including my own performance as a clarinettist, electronic music artist and polymedia artist.

The work *Escapee Gloss*, embodies a response to some old, used clarinets, irreparable and unplayable, and sundry discarded clarinet parts. It integrates these physically and emotionally into experiments with sound, art works, layering of light, reflections and manipulated artwork and music in performance through multi-channel audio and video mixing. The composition combines digitally processed clarinet sounds, live acoustic instruments (flute, Eb/Bb and bass clarinets, double bass and piano), live processed acoustic environmental sounds, props and projections. These components reinvent the disintegrated clarinets through video samples and stills of broken clarinet parts, grey pencil drawings, water footage and Japanese artwork. *Escapee Gloss* explores the possibilities of layering and reflections of light and texture in images and sound and deliberately referencing paint textures in the titles of the movements.

The polymedia processes and multi-art forms cohere in live performance through a deliberate strategy of layering to represent the complexity and depth of the images we see and of the sounds we hear.

1. INTRODUCTION

The initial idea for *Escapee Gloss* came to me when I was given 32 old, used clarinets, irreparable and unplayable, from Kilvington Grammar School in Ormond, Melbourne. Then Wesley College in Prahran said they had a drawer of old clarinet parts that had been sitting there since 1930. These inoperable instruments became the inspiration and foundation of *Escapee Gloss,* a seven-movement sound art composition.

The title *Escapee Gloss* is *A Symphony of Polymedia.* Polymedia is an unusual term and might be seen as a synonym for multimedia, so my use of the term needs explanation. It is, of course, like multimedia in that it refers to using several different media simultaneously. But, whereas multimedia expresses the simple fact of presence, polymedia, to my mind, expresses a sense of unity, too a sense that this mélange of media works towards a common goal and makes up a single coherent work of art. This is why I have also used the word symphony, which implies unity in diversity. Whether one takes the word *symphony* with the Greek translation (*symphonia*), or Mahler's idea about the symphony or that of Sibelius, a symphony is something that melds many diverse voices and many diverse elements into a unified whole. A polymedia symphonist does that and melds many other elements as well. Images, video, graphic scores, extended techniques, fragments of other works, even using the video as a moving score for improvisation, all working together, ideally, to express a single idea, or, better, to express the diversity and multiplicity of a single idea.

My practice is the performer, composer and visual artist creating an the identity of the whole, integrating acoustic sound, live audio mulching and acousmatic sound, all composed for reception via multiple loudspeakers, live video mixing and fixed video during performance.

2. DESCRIPTION

In the context of polymedia, *Escapee Gloss* transmutes the energy produced by music into a visual form, recognisable to the eye and exposing the inner necessity of the artist to explore the values both fragile and strong, common to both sonic and visual worlds. The different sensory worlds created in each movement are intended to encourage the listener to think about how music and visuals complement each other, and gain new strength from their combination.

Each movement of *Escapee Gloss* is named after a veneer of paint and explores avenues of texture, reflection, layers, light, colour, timbre, structure and the interaction of different media and acoustic sound and space. *Shine, Scintillating, Gloss, Pantone, Matte, Sheen, Silk.* Each name evokes a different thought or abstract meaning that I associate with the clarinet and manipulated electronic sounds. The musical/visual idea each veneer evokes is a different texture, literally, as each veneer has its own individual attributes.

Each movement of this composition explores a different disembodiment of the clarinet, from the plain presentation of instrument pieces in *Sheen* and the glimpses of clarinet parts in *Silk* to the kaleidoscope of sound and visuals in *Pantone,* in which the clarinet, sound and visuals both is not recognizable. And, in *Gloss,* for a further extreme, the exploration has been extended into different sound worlds and visual layers using electronics and other instruments.

The main structural technique of this composition is layering—both in the audio and in the video—which quickly and efficiently create a variety of complex situations from a simple pile of unusable clarinets. These clarinets and clarinet parts have evoked many avenues, inspiring me to create sound worlds from the intimate to the extreme. This range is also projected in the visual from literally smashed up clarinets and glass to gorgeous bubbles, fine shadows and lines.

This series of compositions investigates the core of the music: tones, notation, instruments, composition, colour, design, line and texture. The following quote from Harry Partch contextualises how I have aimed to find a tangible meaning in *Escapee Gloss* through all the media presented, how I have taken these abandoned clarinets and transformed them into an exciting and sensitive polymedia work:

> The forms that imagination may devise transform the primitive sound-generation ideas into vehicles for new and exciting adventures, and the act of transforming in itself, like a fire by a stream, is an antidote to this age, a transcendence of its materials. And it is a small reaching back, through many thousands of years, to the first men who wished to find meaning for their lives through art.[1]

The title of each movement of *Escapee Gloss* is the name of a different texture of paint, and the bold colors of red, yellow and black are prominent in the videos. Silver, gold and white paint were used in different layers with the bold colours which transmuted many graduations of light continually transforming each piece of art or stills of the videos. Text prevails throughout this composition in all forms especially in *Silk,* to vocalize thoughts of how and why the music has come about. While the musical composition was usually written first, without the fragments of texts and the visuals of the disembodied clarinet underpinning the whole it is these that provide the concept.

The finished movements involved layering of different sounds and visuals, this layering of different textures in the sounds and light diffusions in the visuals created the pulse of the composition that drives the movement in Escapee Gloss. This keeps the composition moving both audibly and visually, so that no matter how complex or chaotic the combinations, the idea is never lost. The essence is to create a traveling moment with a start, middle and end, with many subtle and boisterous gestures and a sense of surprise and humour.

In performance, the musician plays Bb clarinet and controls the audio samples through the laptop and visual samples through the interactive responses on the video mixer, which is directly influenced by the sonic output of the system.

Gloss, Shine, and Silk involve other performers who are directed through notated/graphic/visual scores and cues, while *Pantone, Sheen* and *Matte* rely on the visual output screened in real time. During the performance, one is continually creating new musical ideas that trigger fresh sonic improvisations that then influence different combinations of visual layering that continually and fluidly change the balance between the visual and audio output. Both are reacting off each other at all times. The visual output is based on an abstract representation of how the music is progressing. The audio is affected by the live video feed, by the lighting and by the amount the performers interact with the footage in real-time. Being a polymedia symphonist is about integrating the sound and image structurally.

All audio processing during performance is created using Audio Mulch, a live audio interface for real-time audio performance.

The audio and visuals are all triggered manually. This is purely by choice, as I treat all the components individually in a polyphonic manner. The natural state for an instrumentalist is to have no attachments to the instrument; so, reaction to musical moments is spontaneous. However adding a laptop provides another line of the control that influences the outcome, resulting in a fluid and reactive performance.

The instrumentation, props and media included are the combination clarinet parts, added contraptions – microphones and cameras, multimedia visual works using parts of the clarinet and related woodwind instruments, improvisations, notated solo clarinet compositions and an electro-acoustic sound art installation. The audio works are inspired by and based on a series of short films.

Examples of electronic plug-ins I have used in the processing of the audio samples include delay, stereo delay, transposition, granulation, grain duration, pan and ring modulation. The process continues with manipulation of files into different layers and different channels, concentrating on microtonal interaction between the samples. A similar process is applied to the visual materials, including analysis of brightness, colour, contrast, duration, speed and complexity. The images

[1]Harry Partch, *Letter to William M. Bowen,* (New York: MS Lingua Press, 1967) 119.

have two categories: graphic-based images and film/still images. The sound and image influences the shape and analysis of each of the works. The audio in the compositions uses a real-time environment of acoustic sound and generative structures.

The imagery in *Escapee Gloss* is all about the dissolve and the transformation of the images. Very little manipulation is involved to fade one image to the next; only the length of the fade is calculated. The layering of the transparencies with luma, chroma, speed (pulse), and cutouts dominate many of the visual samples. The aim is to make the still images move through these effects through different diffusions of light.

3. ESCAPEE GLOSS: A SYMPHONY OF POLYMEDIA IN SEVEN MOVEMENTS

Escapee Gloss composed from 2010-14 will now be discussion of individual movements focussing on the first 3 movements *Shine, Scintillating and Silk* that use interactivity and layering of the audio and visuals with a short description of the last *Matte, Pantone. Sheen* and *Gloss*.

3.1.1. ***Movement 1 – Shine***

For Solo Bb Clarinet, electronics and Visuals

Shine explores a palette of extended clarinet techniques and acousmatic sound, focusing particularly on breath, white noise and timbre. The pitch organization is drawn from a twelve-tone spiral with a converging and diverging whole-tone scale. The clarinet part is also scored in graphic notation and live electronics that indicates the contour and duration for use of the scale effect and sample throughout the piece. The visual elements consist of delicate lines and bubbles crossing between greys, blues and sepia gradations. These intersect with the timbres of the clarinet part, electronic sounds and the angular contours of the musical phrases.

A pulse was created through the gentle movement of bubbles and lines both visually and audibly. My palette of extended clarinet techniques and flourishes focuses here particularly on breath and timbre. This, along with a virtuosic use of quartertones in melodic passages, both romantic and lyrical, aims to create an intensely personal, moving and transfixing bubble of fragility. In performance, one should use the acoustics of the space to create stillness and a sense of anticipation. There is a scene with the translucent surfaces and twisting lines traveling between the different layers.

The layering is replicated in the slow moving pulse of the sepia bubbles that prevails throughout, a kaleidoscope of graphics in sepia and blue gradients that at times dominates the space, appears to be a haze and an irregular pulse over the whole visual experience. The clarinet score and the video are a conversation, creating a layer of sensory completeness in the most intimate way.

The pitch organization of the score is drawn from a twelve-tone spiral with a converging and diverging whole-tone scale taken from Slonimsky's *Thesaurus*[1] *and Melodic Patterns*. The clarinet part is also scored in graphic notation that indicates the contours and durations for use of the scale throughout the piece.

Atmosphere is created through the pulse of the visuals created by subtle swirling with hollow sounds that are notated with big graphic notes in the score. These graphic notations are interpreted by the performer using their intuition and timing so they can take between half a second to four seconds according to the acoustics of the space.

Figure 1. Brigid Burke *Shine*

3.1.2. ***Movement 2 – Scintillating*** *For Electronics and visuals*

Scintillating was inspired by portraits of two clarinetists, one of myself and the other of my sister Grania. The portraits have been placed side by side as the opening visual in the film, as if in conversation (See Figure 2 & 3).

Scintillating began as a fully notated score for Bb clarinet and bass clarinet that was then transformed into an acousmatic composition through filters, fragmentation, rhythmic manipulations, layering and visual elements. The colors and texture are bold and simple as is the relentless pulse, two strong characters, forever voicing their intention, creating at times anguish and stimulation.

These dramatic figures were then transformed into a ghostly presence with intense black ink strokes to create, in the words of Clarence Barlow, "a metamorphosis of dancing objects."[2]

The opening is sombre and atmospheric, but this mood is very shortly interrupted by vibrant and energetic motifs. Syncopation and punctuations of repeated rhythms work against each other and timbral

---

[1] Nicolas Slonimsky, *Thesaurus of Scales and Melodic Patterns* scale, (New York: Charles Scribner's Sons 1947), 123.

[2] Clarence Barlow *Visualising Sound – Integrating sound and visual image as artform from Relationships Acts VIII des travaux 2004-2005 de l'Academie Internationale de Musique Eletroacoustique*, (Bourges: Mnemosyne musique, 2004/2005), 23.

alterations move in and out of sustained tones. These repeated rhythms make an amalgam of tonality and pulse that keeps the work in total harmony with itself.

Scintillating began as a fully notated score for Bb clarinet and bass clarinet that was then transformed into an acousmatic composition through filters, fragmentation, layering and visual elements. The materials that have been employed in the creation of the artworks associated with *Scintillating* are canvas, gold, bronze and silver leaf, resin, impasto, acrylic and house paint, silk screen prints, pen and ink and hand printed paper. The video footage used throughout takes as its starting point the two portraits, which include~~s~~ flames seen through green luscious leaves. The colors used for the portraits are contrasting: red, black and silver in my portrait and yellow, gold and black for bass clarinet. These choices of colors and texture are bold and simple as is the relentless pulse, which operates in *Scintillating.*

The images grow smaller or switch all over the screen, depicting rhythms in the audio. The squares expand, merge with different fields of light. The images float as a field of space.

The video are a selection of parts of the portraits processed digitally, creating layered and refined artworks. The video takes the viewer on a journey with parts of the portraits emerging from each other to result in snapshots traveling over the screen. The original concept in the notated Bb and bass clarinet was to make it seem as though only one person was performing while in fact, the Bb clarinet and bass clarinet are in conversation, almost in unison at times, but as the piece progresses the separation evolves musically as well as visually.

Figure 2. Brigid Burke *Self Portrait*

Figure 3 Brigid Burke *Portrait of Grania*

3.1.3. ***Movement 3 – Silk*** *For Bass/Eb clarinet, bass clarinet/voice, flute/voice, prepared piano, electronics and visuals.*

Silk, is made up of random verbalizations about my art. *Silk* is dark and joyous with unexpected "circus" moments. It contextualizes thoughts into an elaborate score with vibrant interactions of different mediums of sound, performance and visuals.

Silk is a strong, soft and lustrous fabric of melodic lines and text combined with natural colours of red, blue and ochre. The interweaving of the melodic line, textures, timbres and refraction of light creates the feel of smooth textures. The silk screenings of text, clarinet parts and lines superimposed create this depth in the visuals and music as seen in Figure 4.

Silk is divided into two sections, connected by the electronics. Section one (duration 5'30") starts with metallic prepared piano chords followed by the two bass clarinets interplaying with each other with multiphonics and low single notes, with whistle tones from the flute growing in and out of each other. Repeated chords from the piano quickly interrupt the long ambient sounds. Then the action starts with these verbal utterances of fragments of text from both the electronics, flute and bass clarinet performer as they interpret the graphic notation. The sounds from the flautist are spoken into the mouthpiece creating a more intimate breathiness with accented starts and ends to the fragmentation of the words. The bass clarinet on the other hand is uttering words that come across as random moments with guttural utterances, gestures and individual notes with melodic gestural flourishes in the lower part of the instrument. The subtle melodic gestural flourishes of the second bass clarinet add to this excitement.

The extremity of registers from all the wind instruments and electronics with the bass clarinet changing to Eb clarinet at 3min18sec, to add to the structured chaos and extremity of sounds with the continuous metallic and distorted pulse from the piano.

The second section begins with sombre and playful electronic vocal sounds, a long section of almost solo melodic angular phrases from the bass clarinet that are punctuated at times by the fragmented piano rhythms and electronic outbursts.

The electronics are derived from voice utterance and these add another layer of suspense with the long sustained multiphonics, whistle tones and passages of melodic phrases in the bass clarinet. These are in conjunction with the extreme high, low and speech-like abrupt sounds produced from, speech, flute and clarinet sources. All sounds are processed, fragmented, manipulated, and mixed through computer software sound packages. The 'peaks' and subtle layers, and repeated notes in the piano ostinato passages of repeated rhythms and dynamically erratic clusters from the wind instruments give the work an conversational feel.

The prepared piano motives and clarinet extended sounds which included voice and multiphonics were notated and recorded then divided into small samples to create the electronic sounds heard throughout. As the process of creating the score was recorded throughout the rehearsal process, the final score became audible and the video samples of both the still images and video clips were added. The refining the audio electronic music samples was similar to the rehearsal refinement of the notated score and this was done in the multi-channel program *Adobe Audition*.

The way the voice was used is one of the most interesting techniques used in the improvisational section. It adds another dimension, especially on the bass clarinet and flute with acoustic sounds and manipulated and random accented vowels that peak with unexpected utterances, as if each of the instruments are in conversation with each other. This improvisational technique adds to the overall timbre and movement of the work. Another technique used is singing higher melodies and playing lower notes. One hears the breaking down of sounds with fragments and imitations of other instruments which leaves the bass clarinet and flute sounding quite extreme. It can be heard especially in the bass clarinet line of *Sheen* and in *Silk* in which it is based around text that is accented with sampled vocal percussive sounds manipulated live electronically in the computer, prepared piano and verbal utterances from wind instruments using different vowels and vocal sounds while playing the wind instruments.

The text at times evolves into complex paths that are made up of an array of lines and shapes. This is achieved using simple plug-ins and effects which are reapplied to the image. The same process is used in the audio samples from the clarinets, piano and flute. Every still/moving image and audio sample is analysed so the visuals/audio shift, add and subtract from each other. The results are appealing patterns that emerge and continually evolve and grow.

The visual media also included wood, gold, bronze, silver leaf, resin, impasto, house paint, silk screen prints, pen and water color paper, pencil drawings, video footage of which the text was printed into and superimposed over to create these series of images. The superimposed multiple images create a flickering effect throughout *Silk.* The silkscreen prints ghostly register of fussiness suggests the difficulty of pinning down the text more than a traditional typeset of black and white text would have done. The role that these mediums play in communicating the expressive intent of *Silk* is expressed through this technique of layering the silk-screened prints of the text.

The acoustic sounds relate to how they have been digitally processed. The clarinet flutter tongues, sings while playing, speaks and whispers words. The percussive sounds from the prepared piano with repeated chords have been muted by paper inserted across the strings of the piano. The flute whispers, voice fragmentations and whistle tones prevail throughout. Thick woodwind textures with layering of virtuosic melodic and angular phrases are supported by the prevailing pulse from the prepared piano.

The acoustic samples are processed electronically and combined with the samples of the text, sound and sampled video projections. These visual components are then layered and manipulated in the computer. The words are manipulated by the instrumentalists using a notated score. They interpret them by accentuating certain parts of the vowels and words which are graphically pictured. See figure 4 of Page 6 of score.

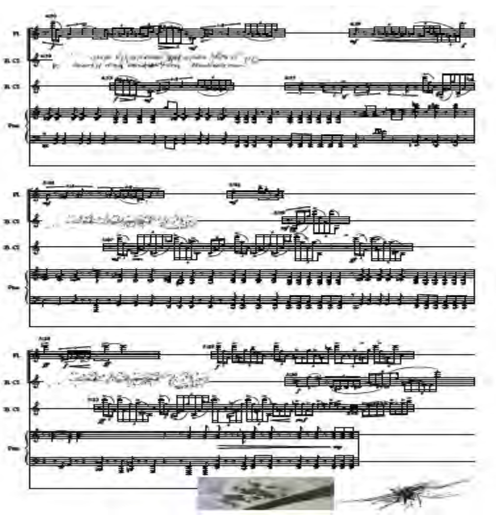



Figure 4. Brigid Burke *Silk* Excerpt from Score Page 6

Each component of sound and visual parts is based on the manipulation of rhythmic pulses created through the selected text that has been spoken, sung and played in different ways by the performers.

The sonic overtones of the woodwinds are explored in detail over the whole work, relating how they are digitally processed from the clarinet flutter tongues, singing while playing, verbal utterances and whispers along with percussive sounds from the prepared piano. The acoustic samples are processed electronically and combined with the pre-recorded samples of the text and sampled video projections. The disintegration of the clarinets is mirrored in the disintegration of text, heightened by the use of polymedia that enables communication on this complex sensory plane.

Timbre is a significant element in *Escapee Gloss,* and this is especially apparent in *Silk.* The use of the prepared piano, where paper is layered on the strings inside the piano, creates quite a unique timbre character. Other timbral motifs in this piece include extended techniques and the melodic fanfare ideas.

The prepared piano creates a prevailing pulse with the repetitive quaver and crotchet chords. These chords were inspired by the repetitive sounds in the piano part of Malcolm Arnold's Sonatina. While Arnold's chords in Sonatina (as seen in figure 3.21) are consonant and form part of the harmonic structure, in *Silk* the chords have been assorted in no cohesive order and play no role within the harmonic structure. By replicating and distorting (due to the paper inserted into the piano) these chords, the timbral effect created is harsh, unresolved and dissonant. The chords are played as block chords or glissando chords and repeated throughout the piece.

The second timbral motif is the use of extended techniques, including multiphonics, overblowing and tongue clicks. However many of the extended techniques, such as voice utterances, hums, screams, whispers, whistle tones and throat sounds, stem from the use of the text. The text plays a significant role throughout and is derived from words and phrases including, "state of mind, a painting, free, chaos, colour, warmth, randomness, stillness, tranquility…" The performer may choose to make sounds using the text by sounding the vowel sounds, a single letter, utter the word or sound fragments of the word. These words or phrases sometimes appear within the graphic notation, improvised sections where the performer is able to draw on their palette of techniques including extended techniques.

Inspired by Arnold's fanfare flourishes in Sonatina and *Fantasy* for Clarinet[1] melodic flourishes appear in the wind parts of *Silk*, however these are employed more for their timbral qualities rather than melodic qualities. Fragments of flourishes appear throughout the piece although they are more prominent in the middle section. It is during this time that the texture thickens as the tutti sections become more frequent. Prior to this build up, the texture is not as thick as instrumental lines interweave with melodic, timbral and rhythmic motifs. These flourishes are disjunct, with complex rhythms. There are fragments of sequences within the motif that are repeated. For example at 4'33'' in the bass clarinet, the triplet ascending idea repeats three times in a short fragmented section. Another example is the ascending first half of the second group of sextuplets at 3'18'' (Bb, G#, E) and the descending second half (high E, A, A), repeated within the fragment at 3'40''.

[1] Malcolm Arnold, *Fantasy for Bb clarinet* (London: Alston & Co. Ltd., Amersham, Bucks, 1966).

The electronic part is a mix of sombre and playful outbursts derived from samples of the text from the voice utterances, and from the percussive piano sounds. The opening begins with a low pulsing drone motif that continues and is fragmented throughout the work. Metallic motifs enter throughout as short fragments or longer passages and the distorted sounds of the prepared piano chords are evident in the second section.

Figure 5. Brigid Burke *Silk*

*3.1.4. Movement **4** – Pantone For Bass clarinet, double bass, electronics and visuals.*

The finished video is the actual score. The performers react to the moving images to create the audio. This is an interactive electronic audio performance that includes spatialized sound and live visual elements. The images of *Pantone* are an ever-dynamic kaleidoscope, interrupted by gentle water moving through the seascape of the glass plate as seen in figure 5. The audio from the bass clarinet, double bass and sampled water electronics makes a subtle counter-point that punctuates and accentuates these delicate moving images, creating depth and interesting nuances.

Figure 6. Brigid Burke *Glass plate with kaleidoscope*

3.1.5. ***Movement 5 – Matte*** *For Visuals and electronics.*

Matte was inspired by the soft environmental soundscapes of rainforests. The clarinet produces wind sounds throughout. The clicking of clarinet keys, percussive cymbal and bowed vibraphone metallic sounds mimic the sounds of rain. This was then visualized in the artwork of *Matte* using round pieces of coloured glass, small broken clarinet parts and both wooden and silver keys, which were moved through water in a glass bowl. These images were combined with a series of blue mixed-media works and coloured pencil drawings on paper formed in the shapes of droplets, a series of squiggle-based mixed media works using acrylic paint, pen and ink and charcoal works on paper which were then photographed as seen in Figure 6.

Figure 7. Brigid Burke *Matte,*

***3.1.6. Movement** 6 – Sheen For Bass clarinet, live electronics/laptop and live feed and pre-recorded video projection.*

Sheen is an improvised quartet for clarinet and real-time audio mulching with live video feed and prerecorded video. *Sheen* has an intrinsic glistening quality and the keys of the clarinet are the focal point for this shining and resplendent movement.

Microphones and a laptop are used to control the synthesis of acoustic material sampled during the performance. The live clarinet interacts with these visual components: a camera filming a bowl of water filled with clarinet parts, another one filming the fingers and keys of the bass clarinet and a prerecorded version of clarinet parts on video that has already been processed. The bubbles and water movement are controlled by an aquatic pump. A pulse is created through such effects as strobes and finger movements from the prerecorded image and live feed of the bass clarinet.

All audio processing is done with Audio Mulch, including the live acoustic bass clarinet manipulated during performance.

Figure 8. Brigid Burke *Sheen, Bent Clarinet*

3.1.7. ***Movement 7 – Gloss – a series of 9 short connected movements.*** *For Electronics, Bb clarinet/bass Clarinet, Eb clarinet/bass clarinet, flute/pic, piano, double bass and visuals.*

Gloss was inspired by a series of nine pen and ink drawings, which were used as graphic notation for the different instrumental combinations and live electronics. The nine graphics were then integrated into the video art work which shows glimpses of the graphics in superimposed layers in conjunction with pencil drawings of squiggles depicting wire as seen in Figure 8 These were then photographed. The depiction of wire glosses the image with a superficial lustre.

Figure 9. Brigid Burke *Gloss 6*

4. CONCLUSION

This paper has examined one major composition, *Escapee Gloss that* represents a map of artistic development and a window into my ongoing artistic practice. As a clarinet performer, I have always enjoyed the immediacy of live performance and the direct engagement with an audience. While Escapee Gloss, I could see infinite possibilities offered through the application of processing to both recorded and live sound.

5. REFERENCES

Books

Partch, Harry. *Genesis of a Music: An Account of a creative work, its roots and its fulfillments.* London: Da Capo, 1949.

Read, Gardener. *Breath and Air sounds: Compendium of Modern Instrumental Techniques.* Connecticut: Greenwood Press, 1993.

Reynolds, Roger. *Notations by John Cage*. New York: Something Else Press, Inc., 1969.

Slonimsky, Nicolas. *Thesaurus of Scales and Melodic Patterns*. New York: Charles Scibner's Sons, 1947.

Weibel, Peter and Gregor Hansen. 2006. "The Development of Light VII From Synesthesia. Synthetic Sounds and Images". in *Light Art from Artificial Light, Light As a Medium in the Art of the 20th And 21st*

Centuries. Hatje Cantz Pub Germany: Hatje Cantz Verlag.

Stockhausen, Karlheinz. 2009. *Cosmic Pulses. Notations 21* Theresa Saucer. New York: Notations 21 Mark Batty.

Xenakis, Iannis. 2010. "How Do You Draw Sound?". *Polytope de Montreal,* edited by, Carey Lovelace,& Sharon Kanach New York: Drawing Centre.

Proceedings paper

Bisig, Daniel and Tatsuo Unemi. 2009 "Swarm Simulations for Dance Performance". *Twelfth Generative Art Conference*, Institute of Computer Music and Sound Technology Zurich University of Arts, Zurich.

Errante, F. Gerard. "The Electric Clarinet, Part I Repertoire." *ClariNetwork* 32, No.2 (2005)
"The Electric Clarinet, Part II Repertoire." *ClariNetwork* 32, No.3 (2005).
"Electro-Acoustic Music for the Clarinet- Part 1" ClariNetwork (Fall 1985).
"The Electric Clarinet Part 1" The Clarinet 32, No 2 (March 2005).
"*The Electric Clarinet Part 11"* The Clarinet 32, No 3 (June 2005).

Kim-Boyle, David. "Extended Non-Linear Applications and Visual Aesthetics of Real-Time Scores". *Proceedings of the International Computer Music Association and the Western Australian Academy of Performing Arts*, Sydney, 2013.

AN INTERACTIVE, MULTIMODAL INSTALLATION USING REAL-TIME SPECTRAL ANALYSIS

Benedict Eris Carey
Hochschule für Musik und Theater, Hamburg

John R. Taylor
Sydney Conservatorium of Music

Damian Barbeler
Sydney Conservatorium of Music

ABSTRACT

This paper describes an interactive sound and light system based on real-time analysis of an augmented musical instrument called the 'motor bow', that when played, makes new, unconventional, often 'noisy' tones. The interactive system comprises a timbre-matching engine that performs real-time analysis of the motor bow performance and algorithmically governed multi-channel diffusion of harmonically similar, but timbrally contrasting material coordinated with visual representations. This interactive installation was designed as part of the Tasmanian International Arts Festival 2015. Of primary importance was that the installation be an extension of the motor bow itself and that it work smoothly in extended performance scenarios. Particular emphasis was placed on extending emergent properties of this interference based relationship between bowed string performance and a weighted, spinning motor placed at the tip of the bow. The entire system can be viewed as an extension of an augmented instrument, one that is multimodal in nature, involving visual (a custom lighting-matrix), audio (acoustic and computer music performance) and tactile (motor bow) elements.

1. INTRODUCTION

1.1. The Acoustic Life of Sheds

The installation was commissioned by the Australian arts and social change company *Big hART* to create an installation in a disused shed in country Tasmania. "Bruce's shed" was due to be demolished, so the authors wanted to make a new performance space to pay tribute to the shed's history. This installation was part of a larger exhibition called "The Acoustic Life of Sheds".

Inspiration for all elements of the installation came rom the character for the site, a disused shed in the far North West of Tasmania. The shed had housed an amateur collection of historical curios in the 60s, 70s, and 80s. The owner was a collector of diverse objects ranging from WWI paraphernalia, to milking equipment, and seashells to indigenous artifacts. The shed was made of wooden paneling that over time had cracked and warped. Light in the shed fell through cracks and pitted windows to a cobblestone floor. All elements of the installation in terms of sound, performance, visuals and the behind the scenes software were influenced by the aesthetics of this building.

The shed, while decaying, had an atmosphere which, far from being dank and sorry, had a kind of quiet, child-like joy: an echo of the enthusiasm of the original man who created the place. Live strings were imagined performing constant dusty energetic patterning sounds of various kinds. Such sounds can be difficult to achieve on a string instrument due to the need to lift the bow regularly to create more bouncing energy. The Motor Bow was invented to solve this problem.

1.2. The Motor Bow

In this installation, the input sources were a conventional viola and cello, played in an improvised style using a modified motor bow (Figure 1) in combination with a wireless microphone. This bow augmentation system was designed and developed by composer Damian Barbeler. The motor bow is a standard string bow with a small electric motor attached at the tip. Electric wires wind down the bow stick to a battery secured under the hand of the player. A moderately weighted front door key is attached to the motor shaft, which when rotating, the uneven weight causes vibration (e.g. overhung rotor unbalance). Since a string bow can be considered a type 1 lever with a dynamic fulcrum (at the string), the radial and axial vibrations coupled with the changes in mechanical advantage as the player draws the bow up and down, produces four distinct bow behaviours, with each behaviour producing a distinct sound. These vibrational behaviours and resulting sounds correspond to four phases of the bow tip draw, relative to the string.

The phases, corresponding vibrational behaviour, and sound are: (1) nearest the hand: jittering, noisy; (2) middle of the bow: tremolo, pitched; (3) top quarter of the bow: small regular ricochets; and staccato

(4) top of the bow: large irregular ricochets, pointillistic, as catalogued by Damian Barbeler. This categorisation method was used in the recording process when building the sample database, so as to provide a diverse range of compositional material for our sample playback system. What was noteworthy about these four phases is how they informed the development of the analysis system. It had to be responsive to both 'noisy' and 'pitched' sound without creating an abundance of random output to ensure accurate, predictable triggering for lighting and sound generate.

Figure 1. The 'motor bow'.

In order to facilitate efficient communication within our team it had to be discussed using terminology familiar to classically trained musicians with varying degrees of familiarity with interaction design terminology.

1.3. System Concept Design

During system design, we used the term "Contrapuntal Interaction" in our own specific way, as short hand for giving the illusion of intelligent reactions from the system. It was felt that the software should "listen" to the player, then respond with sounds which were neither a predictable echo, nor too abstract or random. Rather, a "goldilocks zone" response, reminiscent of the live sound, with some new additional timbral elements. With the motor bow providing a novel set of timbral characteristics, each linked to the physical areas of the bow, groupings of periodicity/aperiodicity, flux, and noise emerged, as the predominant spectromorphological components of the sound (Smalley, 1986; 1997). As the bow enters each of the four physical phases, significant independent variations in amplitude and noise, provided an opportunity to elicit diverse system behaviour over the course of each stroke.

Our solution was a system that identified the spectral flatness of the player's input, and matched the spectral flatness to an opposing spectral flatness value: when the player played a noisy sound, the system would play something with more perceived pitch. In addition, tonality and amplitude information was also extracted from the player's input to refine the system responses, so that they appeared to be the product of a creative mind developing the sound further. It was decided that from an artistic and interaction design perspective, a multimodal system that reacted in a contrapuntal manner would produce a clear relationship between the performer and the audio and lighting system, thus inferring dynamic behaviour. These design choices presented unique challenges for real-time spectral analysis and matching of timbres, while simultaneously maintaining relevant dynamic system behaviour and artistic vision.

2. IMPLEMENTATION

2.1. Timbre-Matching Engine

Designing a real-time 'timbre-matching' engine can be seen as a relatively impressionistic exercise, where the relationship between the input and output streams can be modified depending on the preference of the designer. Our timbre-matching engine contrapuntally triggers spectrally similar material, and operates as a flexible audio-based controller for a custom DMX lighting system. Therefore, from the outset of the prototyping phase, our sample set informed the aesthetic consequences of the system's behavior. Additionally, early design requirements indicated that such a system should be capable of being applied to a variety of audio inputs, in as yet unknown scenarios, so as to be useful for future installation and performance projects without entirely relinquishing control over the selection of the playback material to aleatoric factors.

In collaborating with musicians from a diverse range of specialist areas and educational backgrounds, the common language provided by knowledge of modern musical notation and avant-garde performance techniques by the group became the standard communication convention, and ultimately the framework on which the analysis engine was built. Monophonic pitch tracking was seen as a less useful option in our case where we were interested in triggering material from within our sample database that was differentiated from the live input audio in terms of timbre (noise/tone relationship), yet remained harmonically similar. The presence of multiphonic and chordal viola and cello material was also expected at the audio input and these two considerations impacted on the analysis methods chosen.

It remains a highly processor-intensive task to track timbral changes of input audio in real-time, prohibitively so for use in a system also dealing with multiple modes of output due to the complex calculations involved. In our case, LED mapping was to prove highly demanding in terms of system resources, as was spectral analysis within the Max/MSP environment on the available development hardware (MacBook Pro, Mid 2012, 2.3 GHz Intel Core i7, 8 GB 1600 MHz DDR3, Intel HD Graphics 4000 1024 MB, OS X 10.1 Yosemite). The aim of this analysis engine was therefore to select a series of samples for playback from within a pre-analysed collection, following on from similar efforts by Jehan and Schoner (2001) and Carpentier and Bresson (2010) using a simple analysis model. This was approximated through commonalities of pitch material (extrapolated from FFT data) and a tonality coefficient, which can be measured by calculating the spectral flatness of an input signal (Dubnov, 2006). An overview of the system design is shown in Figure 2.

The analysis component of this system is implemented through manipulation of data derived from the real-time tracking of spectral flatness data, amplitude, and pitch information. It relies on three objects from the Max/MSP environment in particular; *zsa.flatness* and *zsa.freqpeak* (Malt and Jourdan, 2008) and part of the *Max4Live* package *live.slider* (for amplitude tracking) in combination with dynamically updated histograms and databases of 'pre-performance analysis'[1].

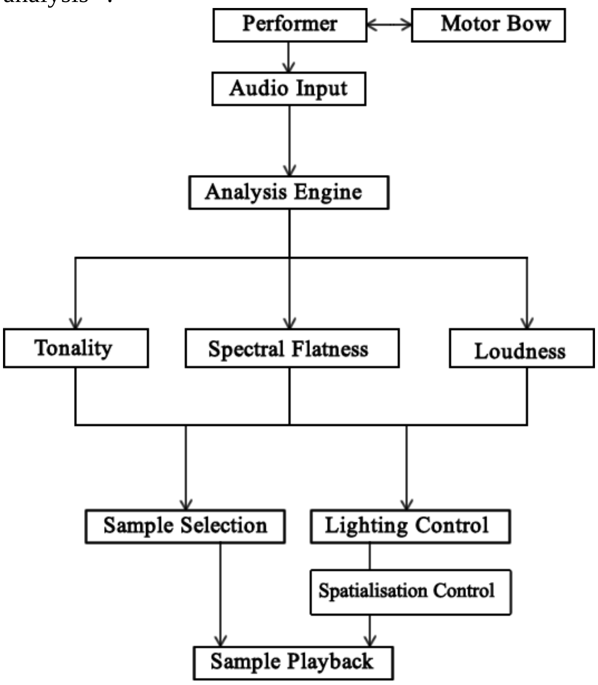


Figure 2. An overview of the analysis system.

Our pre-collected analysis data is based on two databases of motor bow performed viola audio samples recorded at 44.1kHz. The first sample database, comprising of three hundred samples was previously classified into eight groups of increasing spectral flatness, and is sorted and filtered based on the 'Tonality', 'Spectral Flatness' and 'Loudness' values deduced over an analysis period unique to each sample's length. This database was accessed and utilised through the *mxj* Java-framework provided with Max/MSP, again to reduce the overall demand on system resources.

When using the analysis system in a live context, initially, a spectral flatness factor is tracked in real-time over a variable analysis period (referred to as the analysis period throughout the following pages), typically between approximately 1 and 6 seconds, determined by the length of the previously triggered sample. This process determines the most commonly detected 'noise-like to tone-like' ratio of an audio input chunk. High spectral flatness is assumed to correlate to a generally 'less-noisy' signal, and low spectral flatness to a 'more-pitched' (more periodic) signal (Dubnov 2006). This number is stored per window, represented as an integer between 1 and 100, to a dynamically updated collection, which is continuously ranked based on the number of occurrences of each unique spectral flatness value from the most often to least often detected value. This is tracked using the *histo* object in Max/MSP (Figure 3). The amplitude threshold can be modified through the GUI during performance, to filter out lower amplitude frequencies, if desired.

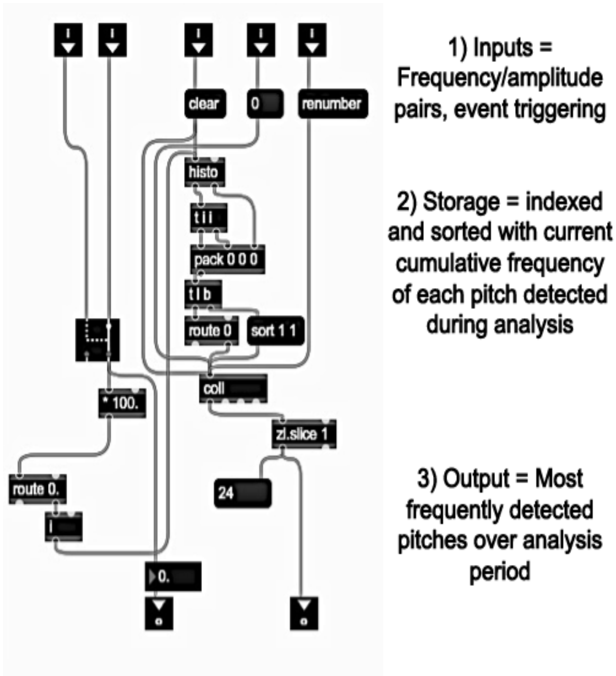


Figure 3. Histograms are used to track the most frequently detected spectral features i.e. pitch/amplitude pairs and current spectral flatness from the same input audio for comparison with the database items in a similar fashion to methods used by Tzanetakis et al. (2003).

Secondly, the most frequently detected spectral flatness value over the analysis period is temporarily stored alongside the relevant string of most significant pitch information, as calculated by the Tonality component of the analysis system. This component outputs a string of 24 floating-point values representing the loudest and most often detected frequencies over the same analysis period. This was implemented through similar use of a histogram function and array as with the Spectral Flatness component. A maximum of 24 unique values was chosen for this string due to the fact that each partial was rounded to the nearest quarter-tone based on a system similar to the one used by Tristan Murail in composing "Gondwana" (Tzanetakis and Cook, 2002). This is intended to allow the performer to interact with the system on a more familiar, music notation related

[1] See Tzanetakis and Cook (2002) and Tzanetakis et al. (2003) for a more detailed explanation of this technique.

level, it is also one way to generalise pitch material for timbre matching.[1] The pitch resolution of the system is therefore 24 discrete divisions per octave or 24-TET (see Figure 4).

The Tonality component differs in that it is based on pairs of floating point values representing pitch and amplitude as tracked by the *zsa.freqpeak* object. This data is derived using a continuously calculated FFT with a window size of 512 samples at a sample rate of 44.1kHz, which also returns 24 pitch and amplitude pairs, per window. Finally, the average amplitude recorded per frequency band is retained in order to further rank the final 24 values, output by the Tonality section of our system from 'loudest' to 'softest'.

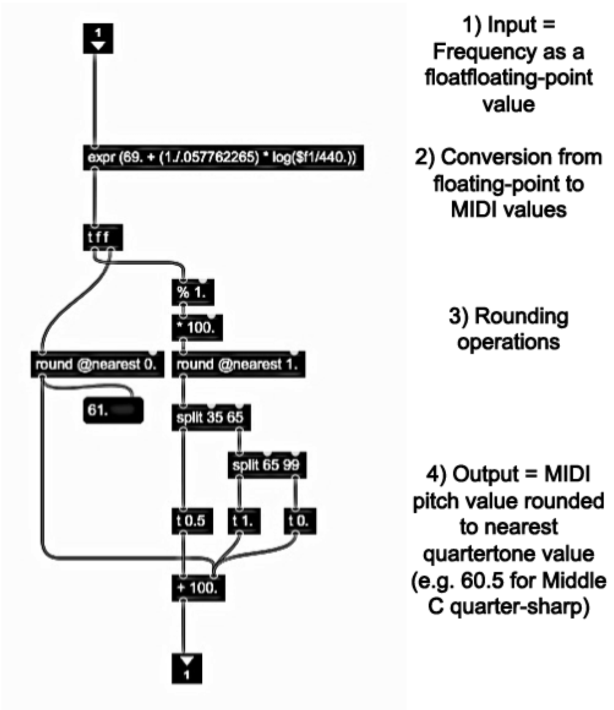


Figure 4. A pitch-quantisation system based on techniques used by Tristan Murail in his preparation of Gondwana. Frequencies are rounded to the nearest quarter-tone (Rose, 1996).

Once the aggregated tonality data and the spectral-flatness coefficient have been determined for the previous analysis period, they are concatenated into a string and stored to an array for retrieval by the Sample Selection component. This component compares the input data with the pre-collected analysis data, and outputs the index of the most similar candidate from the pre-analysed sample set. Internally, the tonality data of each member of the relevant spectral flatness collection is compared to the tonality data derived from the input audio, but only for the most relevant spectral flatness bandwidth group, to determine which sample contains the highest number of coincidence pitches. This data is then passed on to the sample playback engine, which builds musical phrases based on the collected data and synchronously spatialises the audio and lighting across the installation.

2.2. Audio Spatialisation

Spatialisation is achieved using eight individual speakers connected to an external soundcard with eight outputs, placed in the corners and equidistantly of the lighting matrix. The final output of the analysis system is fed into a sample playback probability table and the sample playback engine. In order to maintain spatial coherence between the light and audio, each channel's signal is crossfaded from interpolated coordinates of lighting movement. Since the visualization render was three dimensional, it was possible to change the spatialisation across three axes: X and Y where the audio moves with the spatialisation; and Z, where the spatialisation adds depth of field, and therefore a perceived sense of height, to the installation.

Like the primary sample database, the secondary sample database, comprising of the three hundred motor bow samples, was categorized by spectral flatness values. However, instead of a further categorization by pitch, these samples were categorised by amplitude, and are triggered by the amplitude of the performance input. Spatially, the playback of the secondary samples follows the primary sample set, although owing to the contrapuntal nature of the system, the secondary sample's main function is to add intensity and density in contrast to softer performance input.

In both instances of the playback engine, sample selection is achieved using phrases that comprise of three samples from the collection played with a predefined but adjustable gap of silence between them. The system can be activated using a wireless controller, with three modes of operation: automatic activation and control from the performer, manual activation and control, and interrupt, where certain parameters can be adjusted "on the fly" with a wireless touchscreen controller.

2.3. The Lighting System

The lighting system (Figure 5) is comprised of 1000 addressable LEDs in 10 rows of 100 LEDs, designed to be hung from the ceiling in nine-meter strips, obscured by white balloons of different sizes, in order to diffuse the light. The dimensions of the pixel matrix were determined by the dimensions of the shed. The interactive lighting was created using Max/MSP/Jitter, and then rendered in Jitter Open GL. Syphon Server (Butterworth and Marini 2015) was then used to render the Open GL window inside MadMapper, which was then subsequently transmitted to a PixLite pixel

[1] Murail's method rounds to the nearest ¼ or ½ tone, as opposed to rounding up or down consistently as is the case with the *round* object available in the Max/MSP object suite. Our implementation also has a bias towards 12-tone tempered pitches by using unequal zones where the ¼ tone zone is 30 cents in size while the tempered zone is 70 cents. Due to this generalisation, the system is more reactive. Max 5 and Max 6.1 were used in development and the performance executed in Max 6.1, so issues of compatibility were paramount.

controller connected to the LEDs via the ArtNet protocol.

Figure 5. Lighting System Prototype being tested by Composer Damien Barbeler and Violist, Nicole Forsyth.

2.4. Lighting Interaction Design

The lighting interaction was created in Jitter using the *Boids3d* Max object (Singer et al. 2015) to control Jitter objects that simulate three-dimensional animal flocking in a distributed behavioural model (Reynolds, 1987). Aesthetically, this system sought to impart a sense of depth in the space by having flocks move forward and backward along the length of 100 LEDs. In order to create a consistent visual flocking motion, the attraction point of the Boids was moved back and forth at periodic intervals, although the attraction point of the Boids are constrained to the end of the matrix in the event that the incoming performance audio is both high in amplitude, and high in spectral flatness. In addition, when instructed to play *fff,* the performer triggered colour inversion of the lighting matrix, where the background colour matrices became white, and the Boid black. The Boid behaviour returned to normal once both the amplitude and spectral flatness reduced below their respective thresholds.

Additional visual behaviour includes dynamic changes in the contrast between the "background" LEDs and the Boids, where inactive LEDs have an increase in brightness. This behaviour reduces the brightness of the Boids, reducing their visual impact. In addition, the size and speed of the Boids are also affected by the spectral flatness of the performance input, with higher spectral flatness increasing the size and speed of the Boids.

2.5. Player interaction with the System

The player reacted to this "contrapuntal" system by changing their approach to their performance. This manifested itself in the form of their approach to the inclusion of more wide ranging sounds in terms of the "pure-to-dirty" or pitched vs. noisy spectrum. Although the player was aware that the system reacted to the spectral flatness of their performance, the correlation and translation of such a spectral feature to an identifiable performance metric, led to an exploratory approach to interaction, as the performer sought to illicit more extreme responses by themselves exploring extremes of their own performance norms.

This resulted in more dynamic, contrasting and evocative improvisations by the player. Additionally, the motor bow limited the player's ability to improvise melody. When using the motor bow, the player tended to stray away from melody, instead opting to explore the motor bow equivalent of *klangfarbenmelodie*. The influence of the software in producing more extreme spectral flatness, combined with the player's colouristic, gestural approach, created highly evocative improvised responses. The interaction between the Motor Bow and the system, particularly the system behaviours can conceptually, be considered a proxy score, which in the end produced a repeatable and idiosyncratic work with very specific recognisable features, without the need for detailed notated music.

The sonic result of the final work closely resembles "call and response", where the system generates a contrapuntal response that causes tension as short bursts of noise from the player are contradicted by system generated longer fluttery sounds with more perceived pitch. When the player created more sustained, pitched sounds, the system played shorter noisier sounds.

2.6. Interaction with a Standard Bow

In order to determine the limits of the timbre-matching engine, the performance input was changed to a standard bow. Producing more stable tones with less noise, and with greater relative partials gave rise to interesting and unexpected behaviours. With the motor bow, while there was an element of "call and response" in the system, the contrapuntal nature of the spectral features produced a greater sense of tension in the way the system generated audio material. With the standard bow, the contrapuntal relationship was less evident. This was owing to the selection of samples that were not previously accessible by the system under the timbral constraint of the motor bow. Consequently, the system displayed imitative behavioural responses to the performer. However, owing to the contrapuntal implementation of the visuals (e.g. the size and speed of the Boids) the visual element lacked the kind of variety seen with the motor bow.

3. CONCLUSIONS

The interactive audiovisual installation described in this paper displays very complex, but easily recognizable associative behaviours. The timbre-matching method was chosen owing to the relatively noisy four-phase characteristic of the motor bow. Since the motor bow is a novel extension of the traditional bow, no previous real-time spectral analysis has been undertaken on a bow of this nature, in the context of an interactive installation. The four-phases of the motor bow required constant analysis, with which larger FFT window sizes compromised system resources, since the timbre-matching engine, the DMX light system, and the sound spatialisation were running from within the same software. This window size proved adequate for the desired aesthetic consequences of the system. Additional optimisations included reducing and sorting the returned spectral and pitch values from the analysis, targeting the relevant samples through the comparison engine.

Early prototypes found latency between the Boids and the spatialisation, owing predominantly to the reporting rate of the Boids coordinates, and the pixel rendering. This was because the Jitter window, being at different dimensions (320 x 240) to the pixel matrix (100 x 10), was causing the Boids to stretch and thus producing a sense of latency.

The exhibition "Ten Days on the Island" for the Tasmania International Arts Festival has led the authors to create a complex interactive system that performs real-time analysis on a new type of augmented musical instrument. Owing to the uniqueness of the timbres created by this instrument, a real-time timbre-matching engine was created to make use of the new timbres that proved to be less processor intensive than previous efforts made by the authors with similar software tools.

Further work could include extending the timbre-matching matching engine to other instruments, such as vibraphones, idiophones or membranophones and, as previously described, with a standard viola bow. By focusing on the analysis engine, it may be possible to use a greater percentage of the computational overhead, thus improving the resolution of the FFT windowing size and the overall resolution. In addition, there may be wider applications of real-time analysis, by searching for other spectral features, for example, spectral centroid or bark scale, and could be used in conjunction with other instruments, in addition to more complex feature extraction techniques such as that proposed by Khadevich and Omologo (2013) to potentially cover larger sample databases. A transportable version of this installation is currently being developed, with the lighting matrix housed in an inflatable structure, for demonstration in temporary exhibits.

3.1. Acknowledgments

The authors would like to thank Big hART, The Bundanon Trust, Tasmania International Arts Festival, Synergy, Georg Hajdu for his comments on the pitch quantisation system and Bruce's family for the loan of his shed.

4. REFERENCES

Butterworth, T., and Marini, A. 2015. "Syphon Framework." Available online at syphon.v002.info. Accessed 19 July 2015.

Carpentier, G., and J. Bresson. 2010 "Interacting with symbol, sound, and feature spaces in orchidée, a computer-aided orchestration environment." *Computer Music Journal* 34(1):10-27.

Dubnov, S. 2006. "Spectral Anticipations." *Computer Music Journal* 30(2):63-83.

Jehan, T., and B. Schoner. 2001. "An Audio-Driven Perceptually Meaningful Timbre Synthesizer." In *Proceedings of the International Computer Music* Conference, pp. 381-388.

Khadkevich, M., and M. Omologo. 2013. "Large-Scale Cover Song Identification Using Chord Profiles." In *Proceedings for the International Society for Music Information Retrieva*l, pp. 233-238.

Malt, M., and E. Jourdan. 2008. "Zsa. Descriptors: a library for real-time descriptors analysis." In *Proceedings of Sound and Music Computing (SMC)*, pp.134-137.

Reynolds, C. 1987. "Flocks, herds and schools: A distributed behavioral model." In *Proceedings of ACM SIGGRAPH Computer Graphics,* vol. 21, pp. 25-34.

Rose, F. 1996. "Introduction to the pitch organization of French spectral music." *Perspectives of New Music* 34(2):6-39.

Singer, E., et al. 2015. "Boids for Max." Available online at s373.net/code/. Accessed September 2015.

Smalley, D. 1986. "Spectro-morphology and structuring processes," In S. Emmerson, ed. *The language of electroacoustic music.* Basingstoke: McMillan, pp. 61-93.

Smalley, D. 1997. "Spectromorphology: explaining sound-shapes." *Organised Sound* 2(2):107-126.

Tzanetakis, G., and P. Cook. 2002. "Musical genre classification of audio signals," In *Proceedings of IEEE transactions* on *Speech and Audio Processing,* vol. 10, pp. 293-302.

Tzanetakis, G., Ermolinskyi, A., and P. Cook. 2003. "Pitch histograms in audio and symbolic music information retrieval." *Journal of New Music Research* 32(2) pp. 143-152.

Wechsung, I. 2014. "What Are Multimodal Systems? Why Do They Need Evaluation? - Theoretical Background." In I. Wechsung, ed. *An Evaluation Framework for Multimodal Interaction*, Switzerland: Springer, pp. 7-22.

A FRACTURED EARTH: SONIFICATION AND SPATIALISATION OF SEISMIC DATA WITH MAX/MSP AND SPEAKER.MOTION

Paul Dunham
New Zealand School of Music
Victoria University of Wellington
paul@dunham.co.nz

Bridget Johnson
New Zealand School of Music
Victoria University of Wellington
johnsonbridget.d@gmail.com

ABSTRACT

This paper provides an overview of a Max/MSP patch that sonifies seismic data and the use of spatialisation to explore the resulting soundscape. The background to the development of the patch and spatialisation techniques are discussed and a brief overview of the implementation of the patch is provided with contexts for its use. Spatialisation is explored with the use of a new mechatronic speaker system and its use in a live performance. Finally, the paper considers future development of the system.

1. INTRODUCTION

A Fractured Earth is a Max/MSP based implementation of the sonification of seismic data - the sound of earthquakes. An earthquake is an event that is typically felt rather than heard. We tend to hear the effect of an earthquake rather than the earthquake itself.

This paper presents the motivation behind the project, the process used to capture, store, transform seismic data into sound and spatialisation techniques based on the speaker.motion system. It describes the techniques that have been applied in an interactive based performance to more fully explore the combination of sound and space. Finally, the experiences provide ideas for further development of the Max/MSP patch and how it can be utilised in different performance and installation environments.

2. RELATED WORKS

2.1. Earthquake Sonification

There are numerous and varied examples of the sonification of seismic data (available from the Internet). Examples of such works include r x2 by Dmitry Morozov (Morozov 2015), Tim Prebble (Prebble 2015), Micah Frank (Frank 2015) and SeisSound by Debi Kilb et al (Kilb 2012). These and other examples have ranged from a general interest by people to create a small sonic piece to larger installation pieces. Examples have used a range of sonification techniques from exciting sounding objects in acoustic pieces to synthesised sonic interpretations of seismic data. Some examples have included visual representation of earthquake statistics such as date/time, location and magnitude through static maps or Google Earth. The sonification of seismic data has varied with each person using a different technique to represent their interpretation. This diversity/uniqueness was part of the inspiration for me to explore this area of sound creation.

2.2. Spatialisation Techniques

There is a strong tradition within electroacoustic music of utilising spatialisation techniques to add a further compositional element to new works. Early speaker orchestras such as the Gmebaphone (Clozier, 2001) and the Acousmonium (Desantos, 1997) used variation in the placement and the type of loudspeaker in order to colour the compositions and to more fully utilise the acoustical properties of the performance space.
The advances of Wave Field Synthesis, Ambisonics and Vector-Base-Amplitude Panning have allowed composers to create phantom source positions within the performance space so the audience can also draw meaning from precise spatial positioning and spatial relationships in the piece. There has also been some more recent development in the field of extending the loudspeaker itself. Bahn and Trueman (Trueman, 2000) created the hemispherical speakers as a way to simulate the complex radial patterns created by traditional musical instruments. These speakers also allow electronic musicians to create the same point source localisation when collaborating on stage that would be experienced through localising different sounds to specific instruments in the orchestra.

The research presented in this paper aims to combine areas from all the fields mentioned above.

3. A FRACTURED EARTH

3.1. Motivation

The motivation for this project was to explore the sounds of earthquakes by taking seismic data from an Internet source and transforming this into some form of sonic representation. A number of data types and streams were available so part of the project was deciding on the type of information that was easily accessible.

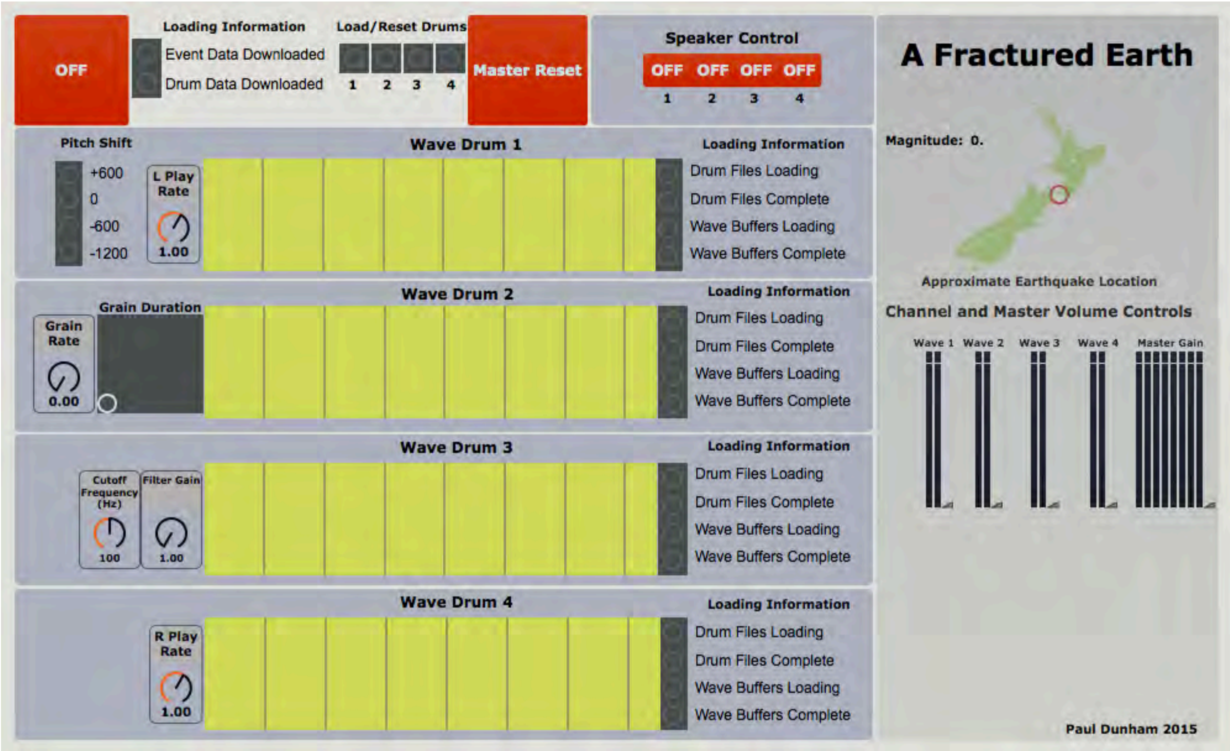


Figure 1. A Fractured Earth User Interface

Living in Middle Earth on top of fractured earth, it seemed appropriate to use a local data source. Data comes from Geonet, a partnership between New Zealand government and research agencies that monitor and research geological hazards and present public information on events.

3.2. Implementation

The original project proposal was to be able to sonify and spatialise seismic data and give another dimension to the reception of this sound. A Fractured Earth, built in Max/MSP, has been built with three design objectives in mind, to be modular, adaptable and extendable. The patch is primarily based on five modules. These are the data gathering, data transformation and storage, sound processing and synthesis modules and audio control.

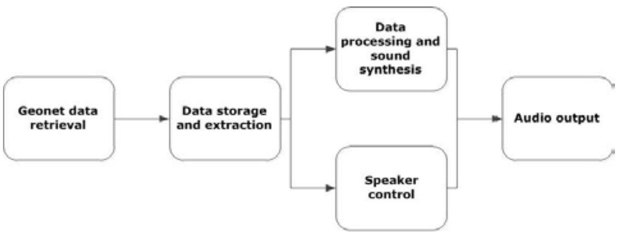


Figure 2. A Fractured Earth Module Overview

3.2.1. Data Retrieval

Data retrieval is a two step process. An http query obtains information on the latest seismic event and downloads the event information in json format. The data is then parsed to extract information from a number of fields. The Public ID is used to compare the current event with the previous one so unnecessary files aren't downloaded, Max/MSP spawns a command line shell[1] to run a script that downloads the data files and stores them on the local device.

3.2.2. Data Transformation and Storage

A shell script queries the list of filenames and this filename information is stored in a coll object for each sonic wave representation (four in this case). A randomly selected file is extracted from the list of filenames and parsed to the data storage and transformation module where it is manipulated to drop unwanted information and to include indexing information then stored.

The extraction process is triggered between two patches that load the seismic data into a wave buffer and represent it as an audio-like waveform. Stored seismic data is parsed to a buffer object for use by the various sound processing and synthesis engines.

1 Bernstein, Jeremy. 'Shell Max/MSP external', in Cycling '74. https://cycling74.com/toolbox/bernstein-shell/#.VhOPaROqqko (6 October,2015)

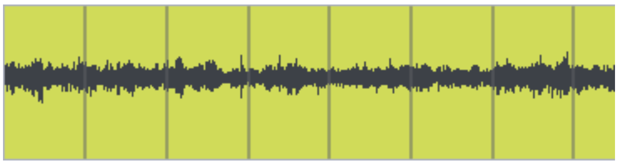


Figure 3. Representation of seismic data as a waveform in Max/MSP

3.2.3. Sound Processing and Synthesis

With the flexibility of Max/MSP, a myriad of different processing effects can be created. One of the hardest things is deciding what to do as a generic processing technique as each set of seismic data could be best suited to different processing techniques and outcomes. To date, the following sound processing and synthesis techniques have been used in this project.

- Time stretching, looping and pitch shifting with interactive playback control.
- Granular synthesis[1] with interactive duration and rate controls.
- Low pass filtering to emphasise the subsonic qualities of the waveform with interactive frequency cut off and filter gain controls.
- Time stretching with interactive playback control

As the system is modular it can easily be adapted to create new processing modules and to select the ones to use based on the performer's criteria.

3.2.4. Audio Control and Spatialisation

There are two components to audio control. The first is simply presenting the audio output. Depending on the output, it is simply a case of selecting the appropriate device and using the individual and master gain controls. The second component is a MIDI control module and is used to control the movement of the speaker.motion system (ref. section 4). This system's flexibility and ease of use means that movement can be fully automated or left to the performer to control. Individual MIDI sequences are sent to each speaker to control its movement. A library of MIDI control parameters can be defined for different types of installation spaces, performance criteria and speaker positions if necessary. These can be easily stored and recalled using Max/MSPs Preset object.

3.3. Use Cases

A Fractured Earth was originally developed as an installation piece where most parameters and controls were automated and the soundscape was explored over a large timeframe. Since using the speaker.motion system, the patch has been adapted to a more performance oriented work.

3.3.1. Performance Use

Performance use of A Fractured Earth allows the performer to explore an environment for a short period of time with a limited data selection. To make the most of this experience, a number of automated controls are removed and the performer can change these as the interaction of sound and space is explored. The performer can change parameters within each sound or change each data file to explore a new waveform. Speaker movement can be independently controlled in real-time as the performer explore the diffused sound field.

3.3.2. Installation Use

The idea of establishing this work as an installation piece was to allow the sonic output to be explored over a large time period as different seismic events occurred. While earthquakes are a relatively regular feature of our landscape, allowing a work to play over an extended period of time would allow a far wider range of soundscapes to be experienced or explored. Placing this into different environments would also allow the interaction of the sound within the space to be explored.

4. SPEAKER.MOTION

The initial idea for spatialisation was to move the sound around a multi-speaker configuration to create a phantom source position corresponding to the earthquake's geographical location. While this may have provided some spatial context, it seemed like static sounds moving around a circle. Enter speaker.motion.

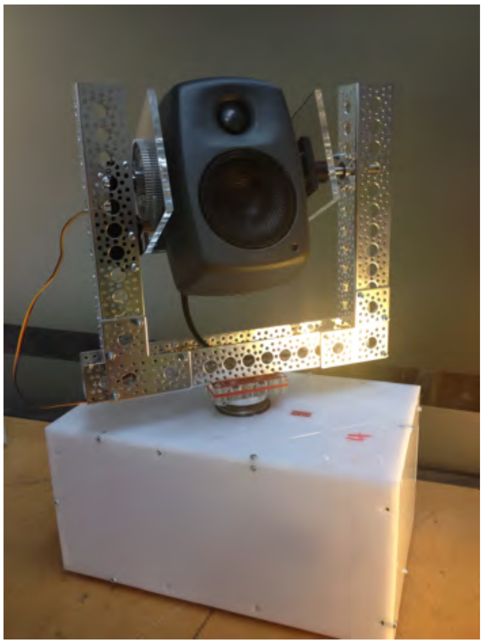


Figure 4. Mechatronic speaker

[1] Hosale, Mark-David. 'grainbuffer~' in Mark-David Hosale http://www.mdhosale.com/grainbuffer/ (6 October 2015)

speaker.motion is a mechatronic speaker system that can be used to manipulate the spatial qualities of electronic music in real-time. Each speaker has the ability to continuously rotate 360º or be accurately positioned within that sphere. In addition, the speaker can be tilted approximately 180º on its vertical axis. Again, this positioning is capable of continuous movement or can be accurately positioned. Control of the speaker system is by MIDI commands. A set of commands have been established to provide various movement commands. These can either be pre-programmed or changed in real-time. These control sequences can be controlled from any MIDI out capable device.

speaker.motion is integral to the sonification because it contributes a new dimension to the sound. speaker.motion allows the sound to be diffused in a space in a number of unique ways. The speaker placement can be determined by the sound being generated and the characteristics of the space the sound is in. The independent movement of the speakers and the control of them allows for a sound or soundscape to be explored in different ways during a performance or installation. Two key benefits of using the speaker system are that the presentation of the sound can be adapted to the type of space it is played in and the use of multichannel audio can further enhance the independent sounds of each speaker.

5. PERFORMANCE WRITE UP

A piece was performed to explore the soundscape created by A Fractured Earth and speaker.motion. The speaker layout is shown in Figure 5.

Figure 5. Speaker Layout for Live Performance Example

Speakers 1 and 4 were configured to rotate 360º and tilt though approximately 90º of their axes. Speakers 2 and 3 were configured to move through 180º facing out into the space and to tilt to pre-programmed positions within a 90º axis at timed intervals. Each sound was loaded individually and each speaker set to move when the sound was fully loaded. While the overall sound density increased, it constantly changed due to the movement of each speaker and the diffusion of the sound. Two aspects during the performance changed the sound that was being explored. First was the ability to change parameters of each sound or to change the sound itself by loading a new data file. Second was the movement of the speakers. The speaker movement had two effects on the sound. First, the collective speaker movement changed over time therefore changing the angles of sound diffusion and reflection. Second, the speaker movement changed the sound's reception by changing the location and strength of the sound. This allowed the listener to hear different aspects of each sound and to hear the change in the overall soundscape.

6. CONCLUSIONS AND FUTURE WORK

The exploration of sound and reception of A Fractured Earth is enhanced with the speaker.motion system. The combination of the diffusion of sound in a space and performance of A Fractured Earth allows the soundscape to be explored in a unique way that can be adapted for different spaces and for the characteristics provided by different seismic information. While the performance aspect of A Fractured Earth is a new idea, the initial performance experience has shown that this is an aspect of the piece that can be developed further.

There are a number of areas that can be considered for future development. One can focus on the installation and performance aspects of the system and where these may overlap and differ in use. Additional sound modules can be developed to offer a range of selectable modules based on the type of installation, performance and/or space. Another area that can be explored is to extend the spatialisation by researching the use of multichannel audio to locate additional sound pieces based on seismic parameters such as location, depth and magnitude. Finally, the performance aspect can be explored further by the addition of a control surface to make controlling speaker movement and sound parameters easier during performance.

7. REFERENCES

Clozier, C. 2001 "The Gmebaphone Concept and Cybernephone Instrument." *Computer Music Journal MIT Press*, 25(4): 81-90.

Desantos, S. Roads, C. and Bayle, F. 1997. "Acousmatic Morphology: An Interview with Francois Bayle." *Computer Music Journal MIT Press*, 21(3): 11-19.

Frank, M. "Tectonic." Available online at http://puremagnetik.com/post/3784034202/sonification-of-earthquakes-off-the-coast-of. Accessed June 2015.

Johnson, B., Norris, M., and Kapur, A. 2014. "Diffusing Diffusion: A History of the Technological Advances in Spatial Performance." *Proceedings of The Combined International Computer Music Conference and Sound and Music Computing Conference* Athens: International Computer Music Association/Sound and Music Computing, pp. 126-132.

Johnson, B. "speaker.motion" Available online at http://bridgetdjohnson.com/speaker-motion. Accessed November 2015.

Kilb, D., Peng, Z., Simpson, D., Michael, A., and Fisher, M. “Listen, Watch, Learn: SeisSound Video Products.” Available online at http://www.seismosoc.org/publications/SRL/SRL_83/srl_83-2_es/. Accessed June 2015.

Morosov, D. “r x2” Available online at http://vtol.cc/filter/works/r-x2. (2 June 2015)

Prebble, T. “Christchurch Quakes – Sonified.” Available online at http://www.musicofsound.co.nz/blog/christchurch-quakes-sonified. Accessed June 2015.

Trueman, D. Bahn, C. and Cook, P. 2000. “Alternative Voices For Electronic Sound: Spherical Speakers and Sensor-Speaker Arrays (SenSAs).” *Journal of The Acoustical Society of America* 2000(1).

THE SONIC BOLLART

Tim Gilchrist
The Perth Artifactory
8/16 Guthrie Street
Osborne Park
Western Australia

Meg Travers
Western Australian Academy of Performing Arts
Edith Cowan University
Perth, Western Australia, Australia

Tara Surry
The Perth Artifactory
8/16 Guthrie Street
Osborne Park
Western Australia

SKoT McDonald
The Perth Artifactory
8/16 Guthrie Street
Osborne Park
Western Australia

Morgan Strong
Project Manager, Digital Strategy
Department of Science, Innovation and Information Technology
Queensland Government

Craig Wale
The Perth Artifactory
8/16 Guthrie Street
Osborne Park
Western Australia

ABSTRACT

BollART is a concept developed during the Perth International Arts Festival's 'art-hack'. It was conceived during the art-hack event, and its purpose is to provide a digital art platform in public spaces. The broad concept - which was rapidly conceived under the constraints of a hack event - was to take a mundane item of urban furniture, and repurpose it as a general-purpose digital artwork platform. The Sonic BollART is a reference implementation/prototype developed as a proof of concept for our idea.

1. INTRODUCTION

The Sonic BollART is an interactive exhibit, conceived and prototyped as part of Australia’s inaugural ‘art hack', Hack The Festival held during the Perth International Arts Festival 2015 (Perth International Arts Festival 2015). The brief for Hack The Festival was to "have 24 hours to create a digitally interactive artwork", followed by a week to turn the idea into a functioning prototype. Over the course of the event, our team went from brainstorming ideas, to a functioning reference implementation, to winning the competition.

Our project was conceived as a general-purpose urban digital art platform, transforming a mundane and practical real world object - the bollard - into a vessel which artists, musicians, students and hackers/makers can use as a virtual stage to contain their digital creations. Reflecting the interests of our group as well as current local arts issues, our prototype was both an electronic musical instrument, and a WiFi dead-drop for distributing local music.

Figure 1 The Sonic BollART Prototype.

2. THE ART HACK

Hack the Festival began with an information session, where ideas could be pitched to other attendees and teams formed. Our team - Team Artifactory - formed around the idea of modifying robot vacuum cleaners to rove autonomously around cultural spaces armed with tablet computers, engaging with the public and provoking them in topical discussions.

The idea of argumentative Roombas was swiftly rejected, as we felt that we wouldn't be able to complete a prototype which sufficiently embodied our ideas, nor deal with potential safety issues[1], under the tight time constraints of the event. Further brainstorming led us to the idea that we could fit a Theremin inside bollard. As these objects are often constructed of metal, the cladding could be connected to the Theremin to form the pitch antenna part of the circuit.

Hack events thrive on rapidly evolving ideas, deciding if they'll fit, if they are achievable, discarding them and changing direction as needs be. Our path from vague idea to Bollard to BollART saw us consider the urban furniture available to us. Bins, lights, fences, rails, as well as existing pieces of urban art were considered as the target of our hack. The bollard was ultimately chosen, as they are commonly deployed in situations where people pass them closely enough to interact.

Over the course of the hack, we developed this basic idea to include a wireless music dead-drop, where local bands could upload and share their music. The dead drop would be disconnected from the Internet, so the public would need to be physically present to be involved.

By the end of the hack day proper, we had the beginnings of a physical object. This was refined over the next few evenings until we had a functioning prototype.

3. BOLLART CONCEPT

We conceived The Sonic BollART as a sound installation, which addressed the problem of accessibility within an urban space to those with visual impairment.

The bollard represents a significant opportunity for public art, and a serious barrier to access. We started by thinking "how can we make a bollard more accessible"?

Bollards do not only restrict access, they create barriers. People with vision impairment cannot find bollards easily. Sound addresses this issue. If a bollard can make sound in relation to proximity it makes it more accessible to everyone. And if we can turn them into a musical instrument, then we make this mundane object both playful and accessible. (Strong 2015)

As a mechanism for disseminating local music, a wireless music dead drop was added. Several members of our team are involved in the local music scene, and there has been much discussion across Perth recently on the lack of music venues willing to take live bands into their entertainment roster - a problem not unique to Western Australia in times where venues and new residential buildings are in close proximity. They also noted that finding new audiences for their work - and engaging them - can be difficult. Having a device where local musicians can upload their music for the public to access it seemed like an ideal second purpose for our project.

Our plans also extended to include "Easter Eggs" - secret functionality often present in software, DVDs and Albums (Wikipedia Contributors 2015). Ideas such as playing particular note sequences on the Theremin unlocked "Disco Mode" light shows and playing tracks from the music dead-drop were raised.

With these core ideas in place, we pitched The BollART as a generic art installation platform, where bollard-laden public spaces could be used as ready-made digital sculpture gardens. By providing basic infrastructure (physical frame and a source of electricity and network access), other artists, musicians, makers and students could use the BollART shells as platforms to create their own digital art.

Just like "cow parade" events involving hundreds of custom painted cows placed in cities across the world, the concept of BollART has similar potential. However, the BollART concept has one fundamental difference: bollards are standard infrastructure in the urban environment, and by converting them into BollART, they are creatively coaxed far beyond its basic raison d'être to delight and engage passersby. Whilst also public art, each BollART is part of the real, functional and normally unsung (and unsinging) working fabric of the City.

4. REFERENCE IMPLEMENTATION

Our prototype consisted of:

- Moog Theremini Theremin controlling addressable RGB LED light strips via an Arduino Microcontroller board;
- Raspberry Pi computer running Linux and the PirateBox web application for the dead-drop
- A laser cut MDF frame for housing the components.

4.1. Physical Construction

We built our bollard chassis using the resources of local Makerspace - The Perth Artifactory[2]. The frame was designed to house and support the various components and power supplies, be simple enough to construct the time we had, and sturdy enough to deploy as a functioning prototype. Once designed, the wooden pieces were laser cut and brought back to the hack venue for final assembly.

[1] Low to the ground, free-roving robots in a public space posing a trip hazard to the unsuspecting public, for example.

[2] http://artifactory.org.au/

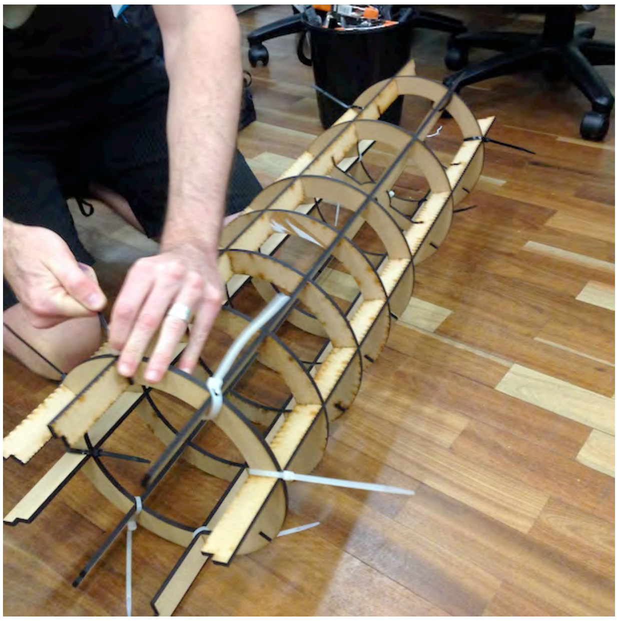


Figure 2 The laser cut BollART Frame under construction.

4.2. Theremin

A Moog Theremini was chosen as the Theremin we would use, as our team had ready access to one and they are more reliable than the kits available from places like Jaycar and other local hobbyist electronics shops. The pitch antenna, which normally stands vertically from the case was removed, and a length of copper pipe was mounted centrally in the upper section of the frame, and wired to the Theremin circuit.

The Theremini was suspended inside the frame by its volume antenna and plastic cable ties. This arrangement proved a less than ideal choice, as both the antennas were in close proximity, making setup of the instrument difficult - little use was in fact made of the volume antenna as ‘players’ would be located too far from the Theremin to make it functional, however the Theremini setup requires that the volume antenna be tuned as well as the now much larger pitch antenna. Performance was deemed adequate for a prototype, so no further modifications were made.

4.3. Wireless music dead-drop

The wireless music dead drop was implemented on a low-cost Raspberry Pi computer running PirateBox[1]. PirateBox is an open-source (free software) Linux distribution, which comes preconfigured with software to run an anonymous file-sharing platform over a private WiFi network.

Although anonymity was not our primary concern, the software was otherwise fit for our purposes. It provided a mechanism for uploading and downloading of files. Android, MacOS and iOS devices were tested, and all were capable of playing music directly from the dead-drop.

Figure 3 Raspberry Pi configured with PirateBox

4.4. Easter Eggs

The “Easter Egg” concept was implemented in our prototype as a software controlled light show using the CV pitch output of the theremin. We decided to make this an obvious function instead of a hidden feature for intrepid users to discover - again due to the tight time constraints of the hack.

Lighting was achieved using addressable LED (Light Emitting Diode) strip. This is a flexible, weatherproof printed circuit board with LED modules and circuitry embedded into it. This strip can be controlled directly from an Arduino microcontroller and several other small, low cost computing platforms.

[1] http://piratebox.cc/

Figure 4 Addressable LED strip installed on the BollART prototype.

The LEDs were arranged as a seven level bar graph pattern, with higher pitches resulting in more LEDs being illuminated. Seven levels were chosen, as that number of LED modules fit comfortably on the bollard frame. The effect we were targeting was that of a 1980's style LED VU meter.

The Moog Theremini provides a configurable CV output. Zero to five volts was chosen for the output range, and fed to an Arduino Uno microcontroller board via one of the analog to digital converter (ADC) pins. This gave us an integer value which varied with pitch between 0 and 1023, for 10 bits of resolution at roughly 4.9mV per increment. Software was written to take the ADC output value, and process it for display on the addressable (software controllable) LED strip.

The 10-bit input values from the theremin were mapped (bucketed) to seven levels, with the remainder used to module the brightness of the final segment in real time. Each segment is fully lit for a value of approximately 146 of the total 1023 range. For example, if the Theremin's pitch output gave a number of 400, the first two segments would be illuminated full brightness, with the third segment at roughly 73% brightness.

LED 1	LED 2	LED 3	LEDs 4..7
0 to 146	147 to 292	293 to 400	N/A
100%	100%	73%	0%

Table 1 Bucketing of an ADC value of 400 for each LED segment.

The net result of this is that when the theremin plays a pitch, the LEDs light up to match.

5. PRODUCTION DESIGN

The ability for the artwork to be deployable in a public space was a key criterion for judging, so we were keen to address this. To expand on our proof of concept, a design for a robust "BollART" platform was produced. A production-ready BollART would be made from concrete, fibreglass and/or metal. 3D renderings were produced for use in our presentation, and a logo was designed.

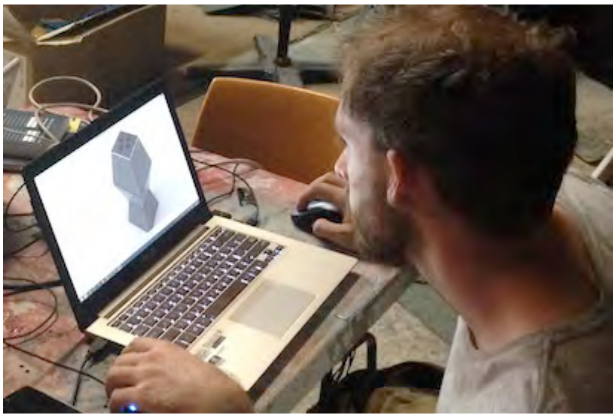


Figure 5 Developing a production design.

Figure 6 The BollART Logo.

6. INSTALLATIONS

The Sonic BollART's initial outing was during the Hack The Festival judging night[1]. Since then, it has been exhibited (or performed) at a number of events including State of the Art Festival, Come Out And Play Festival, The Totally Huge New Music Festival, the Beaufort Street Festival (an annual community street festival held in Mount Lawley, Western Australia), and TEDx Perth 2015.

[1] Footage of our presentation is available online at https://youtu.be/y83jF9PSVx8

Figure 7 The Sonic BollART deployed in the Perth Cultural Centre.

7. THE TEAM

Our Hack the Festival team consisted of: Tim Gilchrist, SKoT McDonald, Morgan Strong, Meg Travers and Craig Wale. We have since been joined by Tara Surry, as we take the project further.

Team members brought a mix of skills to the hack event including: in-depth musical knowledge, engineering, software development, documentation, 3D design, documentation and "making"; all of which contributed to our successful outcome. The maker ethos, in particular, drove our success. A useful description for understanding this ethos is summarised as "The Cult of Done" (Pettis 2009).

8. CONCLUSIONS

The constraints imposed by the art-hack were key in shaping The Sonic BollART. Developing a project during a hack event translates as coming in with broad visions[1], then simultaneously developing concrete ideas; refining them; and ultimately delivering the result. In the case of an art hack, artistic statements and rationalisation are also developed through the same process.

Being forced to rapidly develop our ideas and see them through to an implementation in a short time meant that we were not able to approach it with a clear artistic statement in mind. This had to evolve with the project, and is something which we plan to address as we take our project further.

Applying the maker ethos of "getting it done" did not allow us to overthink the concept, but did encourage us to find creative solutions to obstacles - technical, artistic or otherwise. This provided a degree of focus to our efforts, which proved positive for delivering The Sonic BollART as the winning entry for Hack the Festival 2015.

9. FUTURE WORK

Interest in The Sonic BollART project continues. Our group also has plans to bring the art vessel aspect of the BollART project to production, and have submitted a proposal to the Town of Victoria Park - in response to the Town's call for a long running public digital artwork.

10. REFERENCES

CowParade, 2015 "Our Story @ Cow Parade" Available online at www.cowparade.com/our-story/. Accessed November 2015

Perth International Arts Festival, 2015. "Hack the Festival – Perth Festival" Available online at 2015.perthfestival.com.au/Whats-on-by-Genre/Theatre-and-Circus/Hack-the-Festival. Accessed October 2015.

Pettis B, 2009 "The Cult of Done Manifesto" Available online at www.brepettis.com/blog/2009/3/3/the-cult-of-done-manifesto.html. Accessed October 2015.

Strong M, 2015. "The Sonic BollART" Available online at mstrong.info/blog/sonic-bollart. Accessed October 2015.

Wikipedia contributors, 2015. "Easter egg (media)" Available online at en.wikipedia.org/w/index.php?title=Easter_egg_(media)&oldid=687521570. Accessed November 2015.

[1] at least one per team member.

SCORING FOR GRID CONTROLLERS – PRELIMINARY CONSIDERATIONS OF ACTION, GESTURE, FUNCTIONALITY, AND LIMITATIONS

Vincent Giles
University of Melbourne Faculty of the VCA & MCM

ABSTRACT

In this paper and performance, I aim to demonstrate a set of considerations and influences on a type of score designed for and specific to the 8x8 grid controller commonly associated with software such as Ableton Live.
The grid controller offers a unique method of control that is unlike other MIDI controller systems, and the problem of scoring for it comes primarily from the flexibility of functionality that the controller offers: from mapping to a grid layout of 'clips' within Ableton Live, to conventional MIDI-note systems. The deficiency with this latter mode is the requirement to learn a functionally non-linear map of pitch that is somewhat unintuitive in relation to the traditional staff-based hierarchy.
Through the performance and discussion of the score itself, this paper will show some of the idiosyncratic characteristics of an 8x8 grid controller, possible scoring paradigms, limitations shown thus far and potential limitations moving forward, and further considerations for new methods of working with this grid system.

1. INTRODUCTION

This paper and its accompanying performance outlines the influences and considerations in developing the composition Dead Dirt (Giles 2015) for (augmented) flute and computer performer, utilising an idiosyncratic notational system based on a spatio-temporal design for an 8x8 Novation Launchpad grid controller with Ableton Live.

There is a rich history throughout the 20th and 21st centuries of very specific notation conventions being developed for specific works, or specific players, or specific instruments; a type of notation that is best described as idiosyncratic – a mode peculiar to an individual. Composers have approached the notation for acoustic instruments and electronics in ways that are increasingly individual not only to the composer and the work, but the player for whom it is written and indeed, the instrument itself. The following section shows some examples of instrument-specific notation, and notation for electronics, looking for methods that may be relatable to a method of scoring for a grid controller.

1.1. Overview of Approaches to Notational Movement/Gesture-Mapping

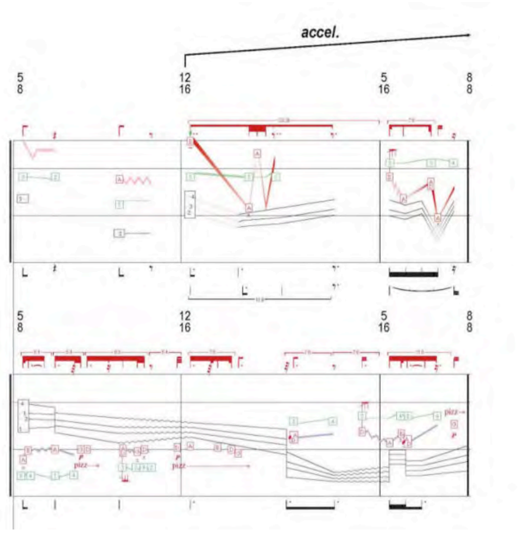


Figure 1. Aaron Cassidy, *Second String Quartet* (2009/10). Mm. 57-59. Violins 1 & 2.

Figure 1 shows a short section of Aaron Cassidy's *Second String Quartet*, in which the composer has utilised a physical mapping of the instruments in question in a type of tablature notation that shows the physical movement necessary in the piece – the performative gesture – rather than the resultant sonic output, which is variable. Cassidy (2013) explains that the staff lines represent, from top to bottom, the bridge of the violin, the end of the fingerboard, the end of the body, and the nut, with bow direction and intensity indicated in green, bow movement across the strings in red, and finger movement across the strings in black. These actions are all rhythmically independent of one another, and the resultant sound is unstable and indeterminate due to variations in bow quality and particularly hand size.

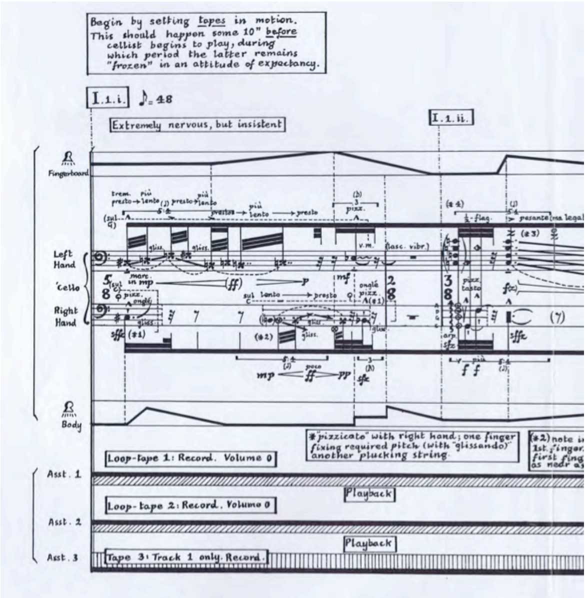


Figure 2. Brian Ferneyhough, *Time and Motion Study II* (1978). Mm. 1-3.

Ferneyhough's *Time and Motion Study II* for solo cellist with amplified cello and eight-speaker sound system (**Figure 2**) illustrates an earlier (than Cassidy) approach to the mapping of physical movement and its influence on sound production. It also shows a type of notation describing the electronic elements of the music, their durations, and the increase and decrease of amplitude over time. Breaking the score down into systems from low to high is the first three tape/computer loops being switched on and recording, with an instruction to enable playback. Notated above and below the cello-proper are the foot-controller envelopes for a microphone placed under the cello's fingerboard, and on the body of it, which the cellist controls (see: (Still 2014) for more detailed information on the technical setup). The cellist needs to control those pedals, play extraordinarily intricate parts on the cello itself, and at some points sing (as outlined in (Iddon 2006)); this bodily interaction with and against the cello is mapped very clearly through the notation and, indeed, the interaction of the body with the instrument is necessary for the piece to be effective as there is a palpable sense of tension in the performance of the work through the cellist's physical actions. Unlike the Cassidy piece in **Figure 1**, this score outlines the *resultant* sound, but the mechanisms for creating that sound are still very much contained within the notation through the decoupling of the right and left hand, as shown in **Figure 2**.

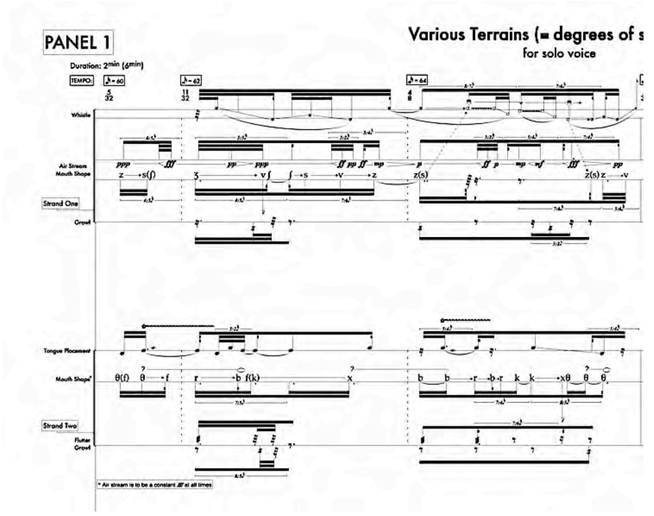


Figure 3. Michael Baldwin, *Various Terains (= degrees of similarity)* (2011). Mm. 1-3, threads 1 & 2.

Various Terrains shows a somewhat similar approach to both of the above examples, with the rhythmic independence of different sound production of a solo voice notated across the independent staves in each system. In practice there are a few possible realisations of the piece, based on the order in which the strands are performed. **Figure 3** only contains the first few bars of the first two strands of the first panel of the score, however it serves to illustrate the importance of mapping the action itself rather than the result of the action not only for clarity, but also for novel and exciting output. The top system (strand 1) contains a whistled sound, the singer's air pressure (and thus dynamic) along with the type of sound to be produced, and finally a growl. It is the interaction and instability of all of these actions that produce the desired effect in performance. It is an interesting example because it combines the extreme decoupling of actions akin to the Cassidy, and hence the instability, with the extreme determinism of the result of that action akin to the Ferneyhough.

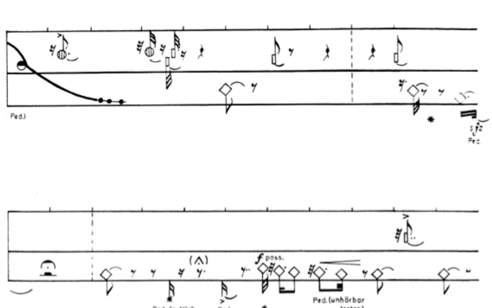


Figure 4. Helmut Lachenmann, *Guero* (1988). Mm. 1-3.

Lachenmann's *Guero*, for solo piano, takes a novel approach to the topography of the instrument, mapping different types of gesture across the surface of the keyboard without depressing the keys themselves, resulting in a percussive sound akin to the folk instrument of the same name. **Figure 4**, this notation is a map of the movement required to produce the desired sound, but unlike the Cassidy the movement is completely deterministic – there are no variables such as hand size – and the resultant sound will be within expected margins of difference from performance to performance. Unlike the Cassidy, the movement does *not* produce an indeterminate sound as a result of the activity. *Guero* is included in this paper as an example that is very close to the desired results of *Dead Dirt* as a scoring paradigm, which will be discussed further below.

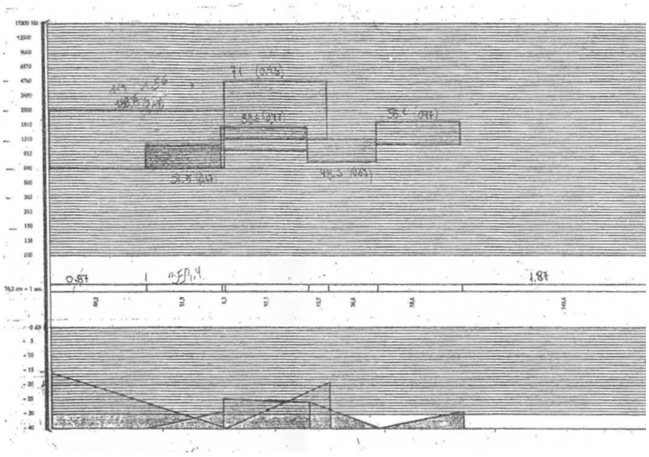


Figure 5. Karlheinz Stockhausen, *Nr. 3 Electronische Studien, Studie II* (1956). p. 1.

The final example of interesting and individualistic approaches to scoring is Stockhausen's *Elektronische Studie II*, which has famously been realised and distributed with MaxMSP for a number of years as a demonstration patch. The excerpt in **Figure 5** is from the original score rather than the Max realisation, and shows Stockhausen's approach to mapping blocks of frequencies (the top part of the system) and their amplitude envelopes (the bottom part of the system) over time (horizontally). While not performative, this is an interesting example of formalism in electronic music, and the gestures created in realisation have an incredible sense of physicality to them, akin to the feeling created by a performance of the Ferneyhough piece in **Figure 2**[1].

As illustrated through these excerpts, in their own right a very small snapshot of the varied approaches to composing for specific instruments and their idiosyncrasies, any approach to notating for a grid controller should be similarly idiosyncratic. Because grid controllers have a non-linear and variable functionality for each of the 64 pads, they provide the ability to reach great distances across octaves without shifting the register of the pads, or without physically reaching long distances. However, the linear visual association with pitch that is common with the piano may also not be the most useful or efficient form of notation for this controller, especially if the computer performer is re-training on this new instrument/controller from another instrument. Not to mention the relevance and directness of expressing something triggered with staffed pitch and whether that has an impact on the inherent semantics of the score. Of course, much of this is dependent upon the outcomes that are desired, and consideration must be made as to whether the grid controller is the best device for the composition at hand.

2. IDIOSYNCRATIC FEATURES OF GRID CONTROLLERS

In light of the above, it seems relevant to work towards unpacking the obvious physical features that distinguish a grid controller from a piano controller, or a set of faders, or an XY touchpad. By making these observations of the physicality of the controller, a composer may start to think about specific notational conventions for that design, in much the same way that Cassidy exploits the physical design of the violin in **Figure 1** and how Lachenmann utilises the physical characteristics of the piano in his composition *Guero* (**Figure 4**) and these instrument characteristics enter into the semantic qualities of the score itself – thus, an idiosyncratic notation for the individual characteristics of the instrument for which the piece is being composed.

2.1. Topographic Layout

The most evident design characteristic of the grid controller is its topographic layout, which relates to the next subheading (N-Dimensional layout), this layout allows for the mapping of triggers (in the form of pitch or any other MIDI message) within a tightly confined physical space. The design of these controllers, while novel in their application, draws upon a history of grid-designs for music dating back to the 1980s in software, and much earlier in human interface design for things like games and even cities (Adeney 2011); all of which demonstrate a degree of non-linear spatial interactivity.

Potentially, every 12 pads from bottom left to top right could represent each of the pitches of the chromatic scale, enabling rapid jumps from one register to another across multiple octaves with only one hand. But more powerfully, it allows the subdivision of the grid into four sets of sixteen pads, enabling simultaneous independent control of different sounds within a DAW – drums, piano, harmonies, and so on, enabling immediate single-performer control of a range of triggered parameters. The flexibility of triggered events within one confined geographic space is immensely powerful, and when combined with the ability to 'shift' the 'register' using the touch of buttons at the top of the controller, this potentially enables very rapid changes between preset groupings of triggers.

[1] This is the opinion of the author at the time of writing, of course other people may not agree.

Figure 6. Novation Launchpad; http://global.novationmusic.com/sites/default/files/styles/cta_scale_640/public/7b-LP-Pan_0.png

There are other instruments that have a topographic layout, all keyboard instruments for example, and even guitars, lutes and similar could be seen as having a topographic layout. None except perhaps lutes offer such a range of pitch in one physical hand space, and none at all offer the range of flexibility for *use* of the space that an 8x8 grid offers. For comparison, a clarinet or a trumpet does not have a topographic layout due to the variability of key or valve depression to achieve linearity of pitch structures.

2.2. N-Dimensional Layout

Strongly linked to the idea of the topographic layout described above is the idea of an n-dimensional layout, although in reality this is limited to only two dimensions (in the Novation Launchpad and Launchpad-S, and three dimensions in the Launchpad-Pro and Ableton Push controller[1]), the X – or horizontal axis – and the Y – or vertical axis. From a conceptual perspective this allows for conceptions of layout in any combination of right to left, left to right, top to bottom or bottom to top, all of which could be interlaced and varied depending on the requirement of the piece and in consideration of the two-dimensional limitation, ignoring for the moment the third dimension of velocity in more sophisticated controllers. This notion of thinking about n-dimensional space has a resonance with Earle Brown's conception of music as activity within a space, and also with the mappings of gesture and movement (both discussed below) evident in the excerpts in the previous section.

Brown (1961) states in the score for *Four Systems*:

> … to have elements exist in space … space as an infinitude of directions from an infinitude of points in space … to work (compositionally and in performance) to right, left, back, forward, up, down, and all points between … the score [being] a picture of this space at one instant, which must always be considered as unreal and/or transitory … a performer must set this all in motion (time), which is to say, realize that it is in motion and step into it … either sit and let it move or move through it at all speeds.

With the additional note: "[coefficient of] intensity and duration [is] space forward and back."

These quotes are accompanied by an image of three-dimensional space represented by a cube with branching dimensions within it, and Brown mentions the instigation of time, both literally and metaphorically, as a fourth dimension within this space. Metaphorically due to the implications of n-dimensional space in music, but literally in that four dimensions consist of the established three: X, Y, and Z, plus *t* for time, that we inhabit as humans.

This n-dimensional space relates strongly to the grid controller, because you could conceive of the 'start' of the device as any point on the X–or–Y-axes and their junctures, which is *not* a characteristic of other controller types, and, linked with the topographic layout described above, is one of the characteristics available for consideration in a scoring system for grid controllers. In the section that follows, I will describe the approach taken in the composition *Dead Dirt* (Giles 2015), the considerations of gesture mapping and movement mapping, the limitations of the controller and the composition itself in relation to a grid controller, and how the inherent characteristic of triggering and pre-designed sound plays a strong role in the compositional layout. I will also describe how considerations of n-dimensional space influenced the layout of the score, particularly as it relates to the time domain.

3. DEAD DIRT: EXAMINATION OF NOTATION SYSTEM FOR GRID CONTROLLER

3.1. Background to the Piece

Dead Dirt was composed for an exhibition opening at the Carlton Connect Initiative in Melbourne, Australia, of the same name. The theme of the exhibition was the notion that our connection with soil is diminishing, and public awareness of the importance of soil on agriculture and life is minimal.

My composition responds to the same thing, utilising data of soil analysis including levels of oxygen, nitrogen, and other compounds as connection points between the nodes in the electronic part as shown in **Figure 7**; it is not a literal transcription of data into music or sound, but an adaptation of the data for artistic purposes, leading to structures that would otherwise not have arisen. I won't detail the process of creation; rather

[1] These dimensions refer to 'musical' dimensions, or the crossover of 'musical' dimensions such as pitch and time, with the physical dimensions of X and Y. Specifically, the Launchpad Pro and Ableton Push both have velocity sensitivity on the pads, combining the two physical dimensions of X and Y with the third 'musical' dimension of amplitude.

I will discuss the mapping of movement in the structure of *Dead Dirt*.

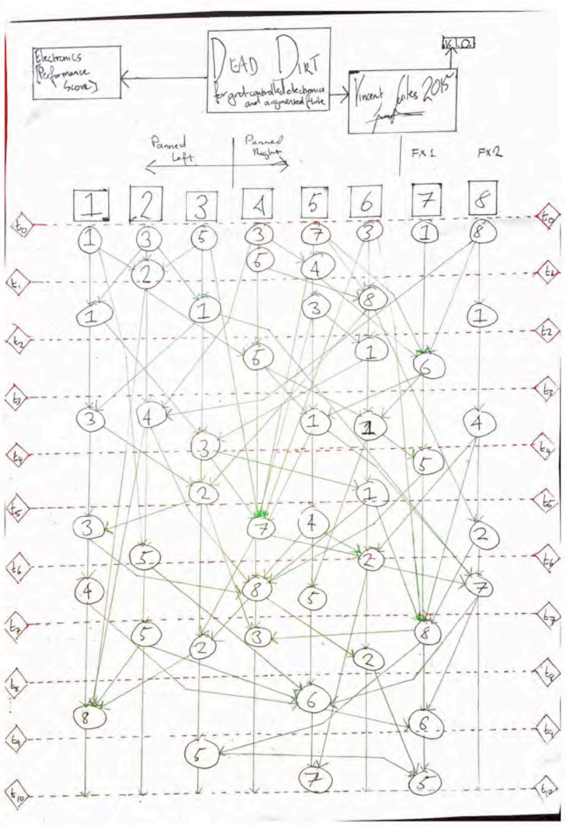


Figure 7. *Dead Dirt*, electronic part (Giles, 2015)

3.2. Mapping Movement

In the construction of the electronic part of *Dead Dirt,* the performance score specifies indeterminate time durations indicated as t_0, t_1, $t_2,\dots,t_9$, t_{10} on the left and right of vertical axis. The first row: [1], [2], [3], … , [8] represents each of the eight columns of pads on the 8x8 grid controller, with a network of 'nodes' moving down the page chronologically. Each node contains a number which represents the row correlating to the column it aligns with; for example, the 3 in the first column says that a little while after t_3, the 3rd row in the first column of pads is to be activated. Whatever event is associated with that pad then becomes active until the movement to the next node.

On most nodes there is a choice of three movements to the next node, which is to be triggered at the relevant point in time which the player should simultaneously deactivate the current node. For example, moving from the same 3 mentioned above to the 8 in column 4 requires the 3 to be deactivated upon activation of the 8. The instruction from the score states: "[o]nce a node is moved on from, the preceding node must be silenced." (Giles 2015)

This 'node'-style score layout can loosely be seen as having been influenced by the Centre for Electronic Media Art music software *Nodal*, an outline of which can be found in (McCormack et al. 2005), and more relevantly, the spatio-temporal concepts that precede Nodal discussed in (McIlwain and Pietsch 1996). This relationship between space and time is evident in my approach the score not only in the electronics, but the flute part too (see **Figure 8**).

Mapping movement across the controller in this way enables a sense of interactivity not only with the controller but also with whatever the pre-determined material happens to be. In treating the triggering of pre-made material as a physical movement; a map of the physical space the controller inhabits that can be negotiated by the performer. This approach echoes those taken by Cassidy and particularly Lachenmann in the excerpts above (**Figure 1** and **Figure 4**) approaching the idiosyncrasies of the instrument (or controller) and trying to approach it for what it is and what it can do, rather than forcing it to a pre-existing idea of notation and the associated hierarchies that come with that, thus mapping a spatial configuration (the X and Y) over time.

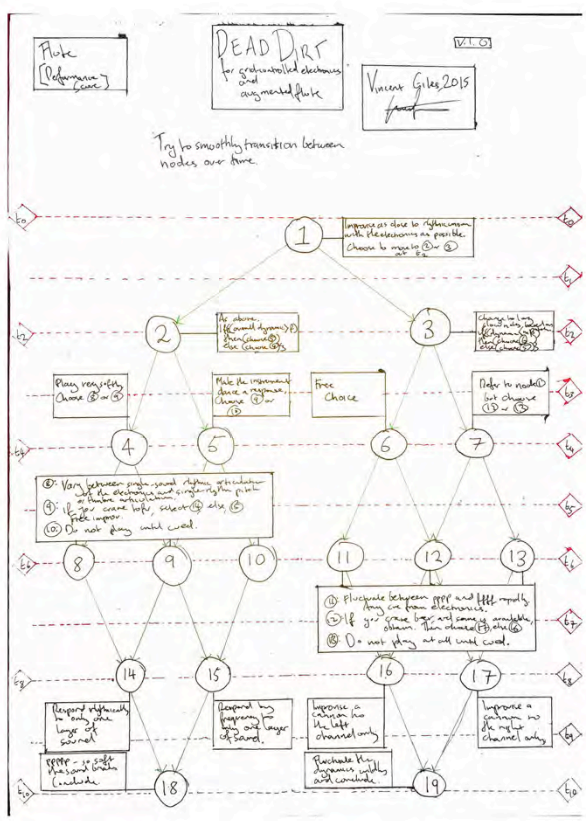


Figure 8. *Dead Dirt*, flute part (Giles 2015)

Figure 9 describes each of the hand-written boxes of instruction in **Figure 8**:

1: Improvise as close to rhythmic unison with the electronics as possible. Choose to move to **2** or **3**.

2: As above.
```
If { (overall dynamic > f)
Then (choose 5)
Else (choose 4)}
```

3: Change to long, slow notes based on frequency/pitch.
```
If { (overall dynamic < mp)
Then (choose 6)
Else (choose 7) }
```

4: Play very softly. Choose **8** or **9**.

5: Make the instrument *dance* a response. Choose **9** or **10**.

6: Free choice.

7: Refer to node **1**, but choose **12** or **13**.

8: Vary between single sound rhythmic articulation of the electronics, and single-rhythm pitch articulation.

9: Free improvisation: if you crave tofu, select **14**, else, **15**.

10: Do not play at all, until cued.

11: Fluctuate between *pppp* and *ffff* rapidly in response to any cue from the electronics.

12: If you crave beer, and some is available, go and obtain, then choose **17**, else choose **16**.

13: Do not play at all, until cued.

14: Respond rhythmically to only one layer of sound.

15: Respond by frequency/timbre to only one layer of sound.

16: Improvise a cannon with the left channel only.

17: Improvise a cannon to the right channel only.

18: *ppppp* – so soft that the sound breaks. Conclude the piece.

19: Fluctuate the dynamics wildly, and conclude the piece.

Figure 9. Clarified node-instructions for flutist in *Dead Dirt*.

3.3. Limitations of Dead Dirt

Dead Dirt shows a number of limitations, the most prominent of which is the limited interactivity *directly with the material at the time of production*, and the second is the requirement for pre-made sound to be linked to the trigger point. In the case of this piece being a semi-improvised work that is functionally for (augmented) flute and electronics, the need for the electronics to be fully interactive for the computer performer was not there. However, in considering the potential of the grid layout as a performative system, then certainly some consideration on how to overcome this limitation is worthwhile.

3.3.1. Triggering, Interactivity, and Pre-composition

The lack of direct interaction with the produced material in favour of simply triggering pre-made material, as mentioned above, is a function of this particular piece. However it's entirely conceivable that a similar movement map could be employed in a directly interactive composition without regard for the MIDI note or any other controller data: in this way taking the notion of mapping movement in a 2-dimensional space as a generator of something directly tangible. Whether or not this is particularly effective technique for the type of controller will need to be explored further, however, the inclusion of the velocity dimension in newer models of the controller may help. In the case of *Dead Dirt* however, the triggering system that is employed is highly effective as the material that is pre-designed and organised within Live's 'session view' (**Figure 10**) is itself highly gestural and interactive. The navigation of the 3-dimensional space-time notation then becomes semi-deterministic (or choose-your-own-adventure) movement through this pre-designed material, giving rise to emergent structures against which the flutist must improvise according to her own set of specific instructions.

The notion of interactivity therefore comes not from an interaction between the computer performer and their own material, but between the flutist and the electronic material; or between human and human-controlled-machine. In a sense, the computer performer therefore becomes a semi-independent agent negotiating a predetermined structure with variables, while the flutist is granted far greater agency and responds not to the computer performer, but to the triggered material – somewhat like a more traditional piece for acoustic instrument and tape/playback. With this design, the pre-designed material needs to be strong and therefore provide something for the flutist to interact with.

Figure 10. Screenshot of the basic 8x8 grid layout used in *Dead Dirt*, without content in columns 7 and 8[1].

With all of this in mind, a brief mention of the importance and considerations when creating the electronic sounds seems pertinent. First, the source for the electronics is sonification of soil-related data, though a different dataset to that used in the development of the score. Describing the full process of sonification falls outside this paper, but once sonified, the wave files form the basis of all the electronic material through manipulations, transcriptions into MIDI and novel sound packs contained within Live, and again transformed into audio for further manipulation. From there, a spatial element is added (as seen at the bottom of **Figure 10**) through panning, and different audio effects are applied to each of the channel strips, so that even if the same material occurs in multiple channels, the output of each will be different. This could easily be adapted to a multi-speaker system beyond the stereo field, in which case the

[1] Due to time restraints during composition and preparation, these two columns were not implemented.

grid controller could take on a much greater sense of spatial configuration and a different scoring paradigm would potentially be needed.

The process of the creation of all material from one source allows a great sense of sonic unity to be achieved, with a great deal of sonic variation and gestural intensity. As the score is navigated by the computer performer, a sense of unfolding and contracting is perceived, and indeed, by looking at the node connections (**Figure 7**) it is possible to see that there are points of convergence upon specific nodes, suggesting a possible expansion and contraction of the sound world that is produced by the computer performer.

4. CONCLUSIONS & FUTURE CONSIDERATIONS

Despite the limitations of direct interactivity discussed above, the composition *Dead Dirt* demonstrates a novel approach to mapping movement across an 8x8 grid controller to create a structured-improvisation drawing on various scoring paradigms and acoustic instruments with pre-composed electronics. As far as the author is aware at the time, due to the grid controller's primary association with more popular styles of music, a scoring paradigm for it has not been explored and yet it is an utterly unique method of controlling a computer system. A spatio-temporal approach to scoring seems the most useful based on the idiosyncratic design of the controller being a 2-dimensional space (X and Y) with the potential to conceive of a beginning at any point on the 64 pads without regard for hierarchical pitch structures. Put simply, treating the controller as an instrument in its own right and looking at its qualities and how best to notate for them.

Further experiments are needed to work towards greater interactivity between computer performer and their material, if that were desired by the composer, and also to look at how these novel scoring paradigms might translate into more popular styles of music, particularly electronic dance music, hip hop, and so on.

To conclude, the scoring system was highly successful for the context it was needed for, and has stimulated further considerations into how to improve on the paradigm, as well as locating many interesting convergences of what might be considered quite disparate practices both within computer music and outside of it.

5. REFERENCES

Adeney, R. 2011. "The HarmonyGrid: music, space and performance in grid music systems." Queensland University of Technology.

Baldwin, M. 2011. Various Terrains (= degrees of similarity). http://michaelbaldwincomposer.wordpress.com

Brown, E. 1961. Folio (1952/53) ; and, 4 systems. New York; London: Associated Music Publishers.

Cassidy, A. 2009/10. Second String Quartet. http://www.aaroncassidy.com

Cassidy, A. 2013. "Constraint Schemata, Multi-axis Movement Modeling, and Unified, Multi-parametric Notation for Strings and Voices." *Search: Journal for New Music and Culture* 10.

Ferneyhough, B. 1978. Time and motion study : 2, solo 'cello and electronics. London: Edition Peters.

Giles, V. 2015. Dead Dirt. Unpublished (lodged with the Australian Music Centre collection; http://www.australianmusiccentre.com).

Iddon, M. 2006. "On the entropy circuit: Brian Ferneyhough's Time and Motion Study II." *Contemporary Music Review* 25 (1-2):93-105.

Lachenmann, H. 1988. Guero : für Klavier = for piano. Wiesbaden: Breitkopf & Hartel.

McCormack, J., A. Dorin, A. Lane, and P. Mcilwain. 2005. "Design Issues in Musical Composition Networks." Australiasian Computer Music Conference.

McIlwain, P., and A. Pietsch. 1996. "Spatio-temporal patterning in computer-generated music: a nodal network approach." Proceedings of the 1996 International Computer Music Conference.

Still, C. 2014. Electric Chair Music. Institute of Musical Research.

Stockhausen, K. 1956. Nr. 3 Elektronische Studien : Studie II. London: Universal Edition.

A DIFFERENT ENGINE

Dr Nigel Helyer
Director, Sonic Objects; Sonic Architecture
Media, Music, Communications and
Cultural Studies, Macquarie University

ABSTRACT

A Different Engine is a kaleidoscopic look at the origins of the digital, driven by pattern making in textiles and music. The paper examines the historical exchange of concepts, images and technologies between East and West via the overland and maritime Silk Trade routes. The paper will reference the importance of the Arabic traditions of Astronomy, Mathematics and Navigation showing how these facilitated this trade, as well as prompting the Renaissance in Europe.

By employing the metaphor of pattern making and the weave of fabrics traded along the silk route the paper will examine the provenance of computer control which can be traced to the early industrial practices textile production, where loom operating instructions were encoded as a series of punch cards, in essence 'digitising' weaving patterns in Jacquard looms.

The virtues of this novel punch card system were not lost on Charles Babbage who adopted them to drive his Difference Engine, from where they were rapidly adapted to automate mechanical music devices, the punch patterns becoming, in effect a form of graphical score capable of sequencing music boxes' barrel organs and later Pianolas.

The Pianola (or Player Piano) was the most sophisticated manifestation of this development and in terms of reproduction quality was far superior to the nascent technologies of audio recording and transcription, such as the Edison Wax Cylinder or disc based Phonography, by virtue of being able to not simply encode musical pitch but also performance characteristics.

Ironically it was the punch card and subsequent punch tape technology that enabled the birth of the modern computer and its entwined history with music.

The first public performance of computer generated music was demonstrated at the Australian Computer Conference in 1951 by a team from CSIRAC (council for Scientific and Industrial Research Automatic Computer) who fed their massive computer with spools of punched paper.

1. THREE THOUGHTS ABOUT CODE; SIGNAL AND NOISE - DISAMBIGUATION

Echo's cries rehearse the utterance of others, departing as counterfeits without significance, returning diminished and disembodied ~ orphaned sounds.

Narcissus swoons as he reaches out to caress the face that has him bewitched. As his fingers glance the perfect image it transforms into an animated mandala, formed of concentric algorithms far more complex than his melancholia.

Smoke curls up from the Beacon Hill, answered in the distance by another and yet still another. A King has died, an Armada has breached the horizon. In every case a presaged message is unleashed ~ only the timing is significant.

Stepping forward through the logic and logistics of the Enlightenment, writing ousts memory and unlike the transient voice, it has the ability to transpose and transport itself ~ it flies and it endures. But like everything it is subject to the 3rd law of thermodynamics, its clarion voice fading with distance. Poured into channels of Copper or sparked into the Æther language swims in an Ocean of Noise in constant fear of corruption, desperate for disambiguation.

The message is quantised, fundamental particles taking the form of semaphore flags, dots and dashes, the texture of Braille. Speech and spelling are rehabilitated as military jingo-jargon, Alpha, Bravo, Charlies.

2. DISTANCE AND TRUTH - COMPRESSION

It is one thing to speak with clarity and be heard over distance (or perhaps even time) but it is another to say a lot and say it fast. The goal of Telematics is to be coherent and robust, economical and fast. Our thoughts, already expressed as serial icons or codes, are now to be compressed into a form that is both necessary and sufficient for the purpose.

Lacking a written language and acknowledging the frailty of memory the ancient Peruvians developed the Quipos, delicate arrays of twisted and knotted threads encoding vital communal information. Marconi abbreviated standard business procedures to save bandwidth and Mawson, who established the first radio communication from Antarctica compressed the limited range of explorer narratives into a code-book, “R-776” meaning I have grown a beard for example.

The compression of meaning and emotion is recirculating today in the form of emoticons, happiness a single condition reduced to a smiling PacMan.

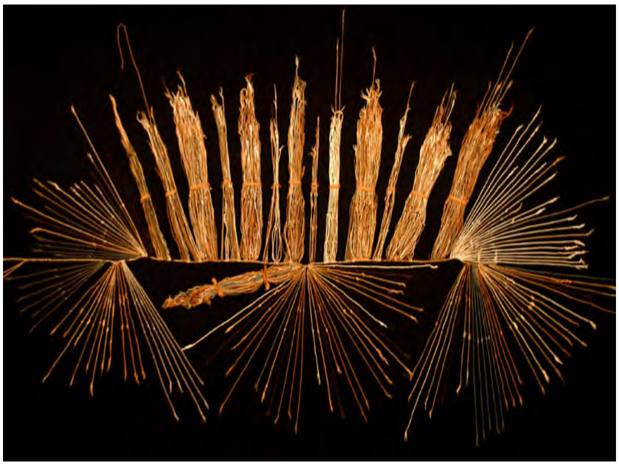


Figure 1. South American Quipos, Anon.

3. KNOWLEDGE IS POWER; ENCRYPTION

Sensitive messages have always been jealously guarded to ensure their privacy and security, never more so than in times of conflict. However, the air, airtime and airspace are open and permeable, available and exposed, as are the transmission technologies which propagate them. The solution to such vulnerability is encryption, the rendering of the unambiguous and compressed into a cryptic form, publicly flaunting itself but impenetrable without a key.

The *Romeo Alpha Foxtrot* (Royal Air Force) held back whilst the Luftwaffe destroyed the city of Coventry in order not to give the Enigma Machine code breaker’s game away, a sacrifice that subsequently sealed the demise of the Afrika Korps through intercepted intelligence, and also initiated horrific reprisal raids on Dreseden.

Across the Atlantic, US army communications were conducted in the unique and modern alphabet of the Cherokee Nation ~ a tongue difficult for the enemy to acquire. In London, at the outbreak of hostilities the BBC panicked over the real possibility of Oxbridge trained upper-class Germans broadcasting ersatz programmes in perfect King’s English.

Their solution was to install Wilfred Pickles as the voice of London Calling. Pickles, a Yorkshireman broke the mould of BBC voice types, with his broad northern accent, impossible even for a Home Counties resident to copy or perhaps understand, in this instance encrypting not the message per se but the vector of delivery.

Figure 2. Wilfred Pickles (BBC Archive).

4. VIGNETTES OF THE UR-DIGITAL; THE SWITCH

Arabia played a fundamental role in connecting the Orient and the Occident via the braided routes of the Silk Route Caravans. These physical pathways also created circuits of transmission of ideas, of technologies and product between East and West. Arabia was also the focal point of trade sea routes, plying between China, India and Europa where the mathematical and astronomical skills of Islamic navigators took pride of place on Chinese Treasure fleets. Viewed in this context Arabia played the role of a gatekeeper of knowledge, rekindling the fires of inventiveness and philosophy in Europa after the deep sleep of the Middle Ages.

The Silk Route becomes a resonant metaphor, a vast entwined network of dusty desert road and sparkling blue sea-ways, the careful and laborious haulage of precious of commodities, all manifestation of luxury embodied as Silk. But what in essence is silk? - a flow of pattern, a flow of intertwined fibres carrying symbolic memories, technological memories and the physical traces of intensive human labour and skill.

5. TEXTO

Silk is text and text is the keeper of memory and knowledge. The Latin to weave is *texo,* more broadly to twine together, to plait, to construct and to build. *Textere* is to compose whilst *textus* is texture. We carry on spinning yarns, knitting our brows, sowing things up and weaving tales. We weave webs of lies and fabrications. The text is a fabric and fabric is a text that travels on the back of a camel across the dunes, and rides the swells in the South China Sea.

6. MEMORY MACHINES

Her hands moves, slowly but inexorably to hover above the keyboard, then gently release to work the keys and the first notes rise from the Organ. She does this perfectly and always with an inscrutable expression. Of course this is the only thing she can do, she is a memory machine, an early automat built between 1768 and 1774 by the Jaquet-Droz family in Neuchâtel. She is remarkable in that she does not mime to an encoded music (i.e. a hidden music box) but she holds in her 'memory' the actions of performance, the movements of head, eyes thorax, arms and digits on the fully functional keyboard, she is a true android performer.

Figure 3. Jaquet-Droz mechanical musician (Ars et Metiers).

7. BASIL BOUCHON

Basil Bouchon was as fascinated by these early clockwork musical automata as he was worried by the fierce competition in the Silk trade, where the genuine article (silk from China) was still cheaper than his Lyon based silk factory could produce.

Indeed his looms were based on the Chinese two person draw bar system which demanded a skilled and fallible "draw-boy" to set the complex patterns. Bouchon, determined to modernise and eliminate his overheads, began to experiment with a mechanism that employed punched paper rolls to control the loom patterns.

Bouchon only ever managed a proof of concept, but his experiments were not in vain. His paper roll system was taken back into the world of music to eventually become the pianola (Player Piano). More importantly they sparked the imagination of another master silk weaver from Lyon, Jean-Marie Jacqard.

8. JEAN-MARIE JACQUARD

Jean-Marie Jaquard developed Bouchon's concept into a robust system that employed chains of perforated cards (one card for each operation of the shuttle - with up to 30,000 individual cards for a single design). This eliminated the second person on the loom (the 'draw boy who's task it was to laboriously select the warp threads) making European looms more competitive with Chinese hand operated machines.

This technology revolutionised weaving but at the same time caused massive social labour disruption (akin to the more recent digital revolution in news and publishing)but most importantly the technological system for encoding information diffused into other areas.

It was not long before the idea of encoding pattern and/or numerical data was taken up by Babbage[1] in England- (1822) and later by Hollerith[2] in the USA (1889 onwards).

[1] Babbage - The Difference Engine (1820's to 1860's) 1822 paper "Note on the application of machinery to the computation of astronomical and mathematical tables" funded by the British Government. Essentially a mechanical (hand cranked computer).

[2] . Hollerith An Electric Tabulating System (1889)Columbia University PhD thesis - data tabulation. Founded the 'Tabulating Machine Company' which later merged with IBM. Basically an electrical data tabulation system employing punch cards built under contract to the US census department - later used by the National Socialists in Germany prior to and during WWII.

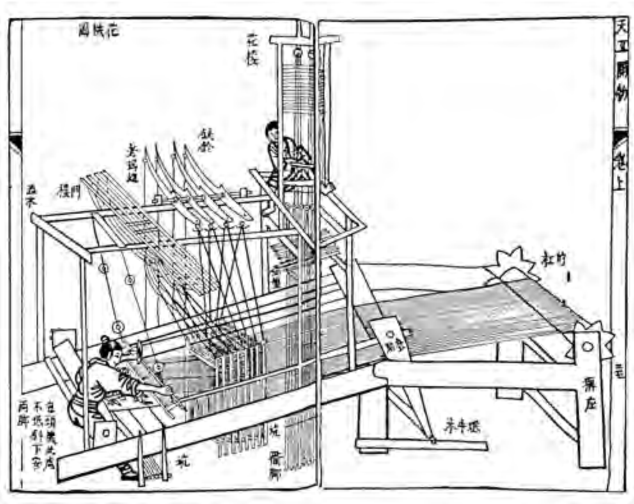


Figure 4. Traditional Chinese drawbar silk loom (Roads to Zanadu).

9. PLAYER PIANOS, MICROPHONES AND GLENN GOULD

But let us not loose the musical thread so quickly. The Player Piano (Pianola) differs from the Piano, from sheet music notation and from a recorded performance in that the smooth rolls of punched paper and the associated pneumatic vascular system, not only encode the music but also the performance values (the precursor of MIDI). Such encoding of content and delivery is reminiscent of Glen Gould, that irascible Canadian virtuoso, who held rehearsal in distain and wasn't too keen on performing in public.

Gould much preferred the recording studio and could be said to have a love affair with the microphone (and of course the subsequent editing process) the technically manipulated results of which, made his recorded performances for the CBS quite outstanding but possibly impossible to match in a live context.

10. THE DARK SIDE

Like Nature, digitisation is indifferent - it can be applied anywhere and anyhow. Hermann Hollerith submitted his PhD thesis "An Electric Tabulating System" to Columbia University in1889.

He was subsequently employed by the US Census Department at the end of the C19th to develop an efficient census evaluation system. In 1896 he founded the 'Tabulating Machine Company' which later merged to become IBM in 1924.

In 1910 the Deutsche Hollerith Maschienen Gesellschaft began operation in Berlin under licence from Hollerith. We can picture Herr Doktor Korherr, loading a batch of census cards into the Hollerith Machine. The year1939 and the National Socialist Census Office has perfected a system to capture the biological make up of each family in the German Reich. We all know the outcome!

Figure 5. poster for the Deutsche Hollerith Maschienen Gesellschaft.

11. A REPRISE

To recap - the provenance of computer control was originally devised as a sequence of punch cards, encoding weaving patterns to operate industrial revolution Jaquard looms.

The virtues of the punch card system were not lost on Charles Babbage who adopted them to drive his Difference Engine and they were rapidly adapted to automate mechanical music devices, punch patterns becoming, in effect, a form of graphical score capable of sequencing music boxes' barrel organs and later Pianolas.

The Pianola, or Player Piano, was the most sophisticated manifestation of this development and in terms of reproduction quality was far superior to the nascent technologies of audio recording and transcription, such as the Edison Wax Cylinder or disc based Phonography, by virtue of being able to encode, not simply musical pitch but also performance characteristics. It was textile and musical patterns manifest as punch card sequences that enabled the birth of the modern computer and entwined its history with music. The first public performance of computer generated music was demonstrated at the Australian Computer Conference in 1951 by a team from CSIRAC (council for Scientific and Industrial Research Automatic Computer).The CSIRAC fed upon spools of paper punch tape.

The world's first true computer, COLOSSUS was installed in 1944 at Bletchley Park for cryptanalysis of the German Geheimschreiber, an in-line cipher teletype machine. This was followed closely by ENIAC, a US military, numerical integrator and computer used to calculate ballistics and the Atom Bomb.

But on the side of light, CSIRAC 1951[3], the first public performance of Computer generate music, Sydney.

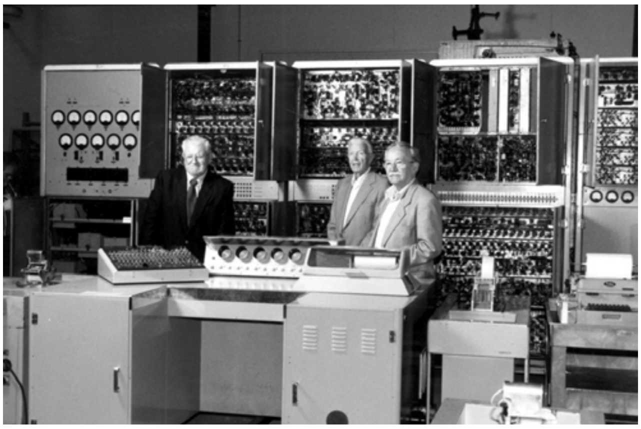


Figure 6. CSIRAC in Sydney 1951 (CSIRO).

[3] see - Paul Doornbusch. http://www.doornbusch.net/CSIRAC/

12. SOME THOUGHTS ABOUT CODE AND LIFE; JOE DAVIS AND MALUS ECCLESIA NIGEL HELYER AND GENEMUSIK.

As an end-piece let us consider that other revolution in encoding discovered during the twentieth century, DNA.

In 2003 artists Joe Davis and Nigel Helyer exchanged a series of ideas for encrypting information in DNA. Davis pursuing textual codes and Helyer musical.

At that time Helyer, working under the aegis SymbioticA, collaborated with the School of Agricultural Science (University of Western Australia) to develop a proof of concept designed to translate music into DNA which when inserted into Bacteria was able to be re-mixed and subsequently extracted and de-coded into novel musical forms.

Fast-forward eleven years and Davis is en-train of realising his Malus Ecclesia[4] project at the Harvard Medical School. Davis plans to transpose the fount of all human knowledge, Wikipedia (sic) within the junk DNA of an ancient strain of Apple. Malus in Latin represents both Apple and Evil (whereas Ecclesia refers to Church - and pays an Homage to George Church the Harvard Professor with whom Davis is working).

This reprise of the Garden of Eden scenario, Davis will ultimately fill a grove of grafted apple trees which will presumably contain all branches of Knowledge. However the apples may be covered by an indictment on consumption, this time not by Jehovah but by the US food and drug administration!

Figure 7. Dr's Helyer and Albertyn at the UFS Microbiology Lab, Bloemfontien SA with GeneMusik bacterial cultures.

[4] see http://www.newyorker.com/tech/elements/object-of-interest-the-twice-forbidden-fruit

In a similar manner Helyer has nurtured his interest in the parallelism between Genes, Memes and Musical Notation as mnemonic structures capable of evolution and the embodiment of memory.

In 2014 GeneMusiK[5] rides again to create a re-mix of cultural, social and biological pathways. Working in South Africa with indigenous musicians GeneMusiK hybridises local ethnic music with the epitome of the western musical tradition, the string Quartet via the transformations of musical and genetic codes within Bacterial cultures. Musical patterns of the indigenous San peoples are genetically transformed to infiltrate the formal notation and performance values of western art music.

Figure 8. Drawing of a San Bushman playing the Gorah with musical annotation, from Travels in the Interior of Southern Africa, Burchill.

[5] see - http://www.sonicobjects.com/index.php/projects/more/eine_kleine_genemusik

13. A THOUGHT ABOUT SOUND AND LIFE; UNDER THE ICECAP.

And a final thought about data. Under the IceCap[6] is a long-term collaboration between Artist Nigel Helyer and Marine Scientist Dr Mary-Anne Lea at the Institute for Marine and Antarctic Studies, University of Tasmania (Hobart).

In a nut-shell the project team renders complex environmental bio-logging data-sets collected by Southern Elephant Seals[7] on their extended under-ice dives and long open ocean transits into 4D cartographic animations, sonifications and graphical music scores, which are used to generate live public performances.

Our decision to interpret environmental data via an aural process is based upon a hunch that musicians have the best pattern-recognition 'wet-ware' around and that our aural sensibility is in fact more finely tuned to detect variations in pattern and recognise subliminal differences, than our visual sense.

Figure 9. Elephant Seals with BioLogging device, IMAS/UTAS.

The byeline for the Institute of Marine and Antarctic Studies is *Turning Nature into Knowledge*. The Under the IceCap project supplies a second line *Turning Knowledge into Culture* encapsulating a powerful Art and Science synthesis. The primary aim of the project is to produce creative work which is compelling and affective but which is simultaneously a work of

[6] see - http://www.sonicobjects.com/index.php/projects/more/bio_logging

[7] Southern Antarctic seals can dive to 2000 metres, stay down for 2 hours and make ocean transits of many thousand Kilometres.

scientific utility - hopefully tapping into both sides of the brain! The key focus is to illuminate the relationship of the environmental knowledge generated from Antarctic bio-logging data with the Anthropogenic changes in the biosphere.

Figure 10. Under the Icecap live performance with data-visualisation screen shots. Helyer.

14. REFERENCES

Babbage, C. 1994. *Passages from the Life of a Philosopher*. New Brunswick, NJ: Rutgers University Press.

Swade, D. 1991. *Charles Babbage and his Calculating Engines*. London: Science Museum.

Fava-Verde, J-F. 2011. *Silk and Innovation: The Jacquard Loom in the Age of the Industrial Revolution*. Leeds: Histancia.

Doornbusch, P. 2005. *The Music of CSIRAC: Australia's First Computer Music*. Available online at www.doornbusch.net/CSIRAC/.

Austrian, G.D. 1982. *Herman Hollerith: The Forgotten Giant of Information Processing*. New York: Columbia.

WHO WANTS A TRAUTONIUM MUST BUILD ONE

Meg Travers
Western Australian Academy of Performing Arts
Edith Cowan University
Perth, Western Australia, Australia

A/Prof Cat Hope
Western Australian Academy of Performing Arts
Edith Cowan University
Perth, Western Australia, Australia

ABSTRACT

The man most associated with the development of the Trautonium, Oskar Sala, stated “Wer ein Trautonium will, muss sich eins bauen” – “Who wants a Trautonium, must build one”. With only a limited number of the instruments ever made, today if you wish to learn the instrument or perform the works composed for it, first you must build your own. This paper looks at the beginning of the journey to research and build a modern Trautonium, and the implications for the preservation of non-commercial electronic musical instruments to allow them to remain in active use.

Keywords: Trautonium, History of electroacoustic music, Music information retrieval, makers, archiving, preservation, modular synthesizer, musicology.

1. INTRODUCTION

This project about the 21st century recreation of an example early electronic musical instrument, utilising modern electronics – the Trautonium. Whilst the scores and recordings of the instrument are an essential part of assessing the success of this recreation, they are artefacts whose preservation is well understood and undertaken. The instruments themselves, remain under preservation and on display in various museums, however this circumstance does not lend itself to ongoing performances on the instruments, nor in the future development of it.

2. PRESERVATION

Preserving the ability to re-perform works, and not simply the output of performance (such as scores and recordings), is the major part of the archiving, curation and preservation of the performing arts (Giaretta 2011). This is a challenge in the re-performance of works for the Trautonium.

For the composers of the 20th century, the issues of preservation are compounded by their temporal location prior to the digital revolution. These musicians used analogue and early digital electronic devices as part of their sonic palette, and the archiving of these objects is more akin to museology than musicology. Much of the work already done on the preservation of 20th century eletronic music, such as the MUSTICA project (Bachimont, Blanchette et al. 2003), is principally related to the preservation of software, rather than hardware sound sources, digital data rather than analogue.

Since the death of Sala, the Trautonium exists as a small number of bespoke instruments in museums in Germany, exhibited as specimens in the museum. Musicologists, music creators and performers are effectively prohibited from using the instrument as it was intended – for the performance of electronic music. Whilst museums are traditionally seen as places that “acquire, conserve, research, communicate and exhibit the tangible and intangible heritage of humanity” (International Council of Museums), they are generally principally interested in the conservation of the item, and not in its continued usability. The Deutsches Museum website speaks of the need to document and catalogue the material left behind by Sala, and only undertake preservation activities on his recordings and the musical instruments (Deutsche Museum a).

These preservation activities have been undertaken by German engineers with some interest in electronic music, and whilst they have not improved access to the original instrument, they have studied the technology used in the instrument and created variations to bring the instrument into the 21st century. Oskar Sala knew his instrument so well that the controls of the sound generation section had no labels on them to guide a player in how the instruments sounds are created, or what circuits Sala is tweaking in his performances. It is only through the investigative work of these engineers that there is now an understanding of how the instrument worked, and how it can be rebuilt.

3. HISTORY

The Trautonium was first developed in 1929 in Germany by Friederich Trautwein and Oskar Sala. Sala, who was the instruments principal performer, passed away in 2002 without having trained anyone in performing on the instrument, and leaving behind seventy shelf meters of notes and recordings of pieces composed for the Trautonium and the instruments themselves in various museums in Germany. However for all intents and purposes, the opportunity for people to perform these works on the instrument they were written for has been closed off. Whilst many commercial units, electronics kits and software instruments exist to replicate other early electronic instruments such as the Theremin, the work that has been conducted to replicate the Trautonium has been highly specialised and almost entirely centred on museum display, yet the Trautonium may have as large a repertoire as the Theremin available to it.

The Trautonium gave Germany an early entry to the development of electronic musical instruments, following immediately after the development of the Theremin in Russia in 1927, and the Ondes Martenot in France in 1928.

The Deutsche Museum details the variations of the instrument were produced for or by Oskar Sala (Deutsche Museum b).

- The first RVS Trautonium from 1929 is located at the Deutsche Museum in Munich;
- The Rundfunk-Trautonium developed for a German radio station in 1935 has decayed or been destroyed;
- The Konzert-Trautonium developed in 1938 was the first instrument used by Sala for concerts, and is currently located at the Deutsche Museum in Munich;
- The Mixtur-Trautonium was developed by Sala after World War 2 which utilised a new form of sound synthesis previously only theorised, subharmonics, for which he received patents (Sala 1956) (since expired) in the USA, Germany and France. This instrument is located at the Deutsche Museum in Bonn;
- The Mixtur-Trautonium nach Oskar Sala (also known as the Halbleiter Trautonium) was developed for Sala by academic staff and students from the University of applied sciences of the Deutsche Bundespost Berlin from 1980-1988. This variation utilised microelectronics and semiconductors. This instrument was the principle version Sala used until his death in 2002. It is currently stored with the Berlin Musical Instrument Museum.

Additionally, between 1933 and 1935, Telefunken manufactured less than 100 “Volkstrautoniums” based on the original version of the Trautonium, but the high price (roughly two and a half months of the average salary) did not encourage people to purchase the instrument and it was discontinued after only thirteen were sold (120years.net) and of those, less than ten are believed to still exist, three of which are with the Deutsche Museum in Munich.

Unlike the Theremin and Ondes Martenot, the Trautonium is rarely used in current music practice, either in art music, popular music or film soundtracks for which it became principally known.

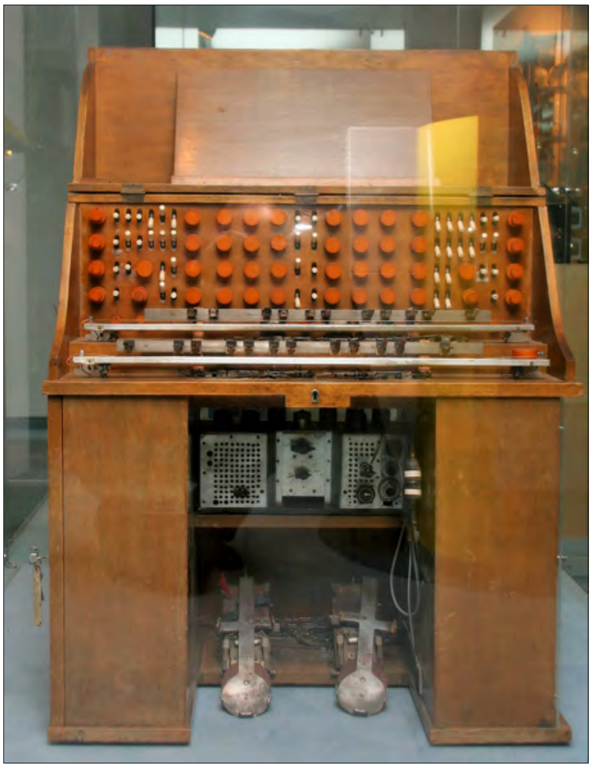


Figure 1. Mixtur-Trautonium on display in Deutsche Museum, in Bonn (photo by MatthiasKabel, Wikipedia). Note the lack of labeling on the controls.

4. LATER VERSIONS

The Trautonium has not been completely ignored by people wishing to recreate the instrument, and a small number of versions have already been developed.

4.1. Replica Mixtur-Trautonium (Vienna)

Doepfer MusikElektronik were contracted by the Technishes Museum in Vienna to construct the sound generation electronics of a replica Mixtur-Trautonium for display in the museum. Dieter Doepfer was able to consult with Oskar Sala in the development of this instrument, and this led to his company creating several synthesizer modules to replicate the sounds of the Mixtur-Trautonium, the prototype modules being installed in that instrument. (Doepfer MusikElectronic GMBH a).

4.2. Trautoniks

Trautoniks are a German company who have the highly specialised business of building faithful reproductions of the Volkstrautonium. Their initial work with the Trautonium was in the development of the electronics for the Mixtur-Trautonium nach Oskar Sala, and providing repairs to this instrument for Oskar Sala in 1990-1995 (Trautoniks 2012). Since then they have built one clone of this instrument for the Berlin Government, and three Volkstrautoniums for Peter Pichler who has recently started performing on these instruments in Germany.

Sala’s involvement with these later projects as well as his use of the Mixtur-Trautonium nach Oskar Sala, demonstrations that he was interested in how the Trautonium may develop using advances in technology. However, given all of these changes to the instrument over time raises the question of what is quintessentially, the Trautonium?

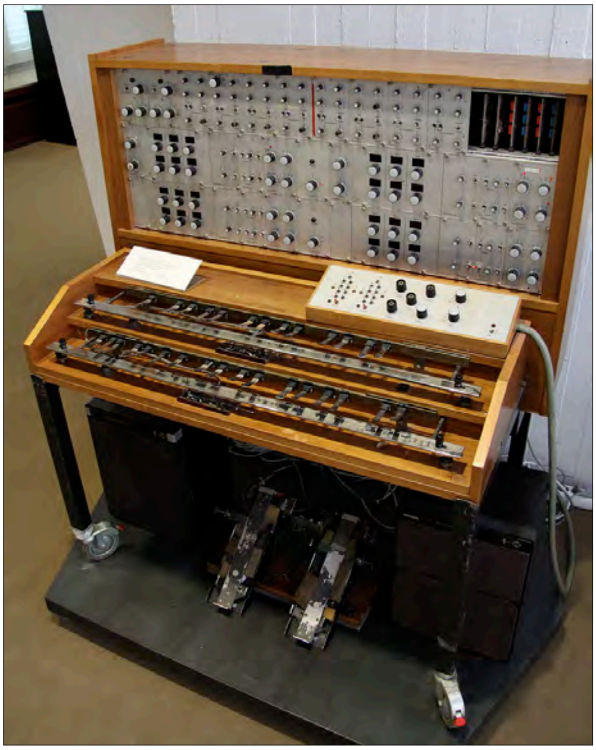


Figure 2. The Mixtur-Trautonium nach Oskar Sala, a much modernized version of the instrument created between 1980-1988.

5. SYNTHESIS AND SOUNDS

Unlike the Theremin and the Ondes Martenot which produced sine waves through heterodyning, the Trautonium generated a sawtooth wave that could be passed through a set of filters (a first example of what is now known as Subtractive Synthesis) and demonstrated Trautwein's theories which were that a sound is not so much characterized by its fundamental frequency, as by the frequency and dampening of the secondary vibrations setup by it which he referred to as formants (Kock 2012).

Apart from this new process for synthesising interesting sounds, the method of playing the Trautonium was also unique, with the performer pressing a wire onto a metal plate (known as the "manual") - an action that is somewhere between playing keyboard and stringed instruments - thus changing the resistance of the circuit and altering the frequency of the sound produced. This method allowed not only the glissandi type of playing that the Theremin and Ondes Martenot produced, but also the ability to play individual notes quite percussively and precisely.

As well as developing subharmonic synthesis in the Mixtur-Trautonium, in 1965, Sala commissioned the Berlin based measuring instruments company, Kamphausen GMBH, to develop for him a Frequency Shifting device based on tube technology. This unit was a one-off, bespoke device and became Sala's "musical calling card" (Trautonist 2012). At this time, devices of this type were rarely seen in the field of music, coming to prominence in the 1970's when commercial units were developed specifically for musical use by Bode and Moog. Developed initially to combat feedback in communication and amplification systems, Frequency Shifters work by moving all audio frequencies by a specific frequency amount (for example all are moved by 100Hz). By comparison, Pitch Shifters work in a more "musical" way by moving the frequencies by intervals (eg: all frequencies are doubled). When using a Frequency Shifter, the overtones in the source sound will become increasingly distorted as the shift is increased, and it is this sound that Sala made extensive use of in his performance and compositions.

6. CHOOSING A NEW TRAUTONIUM

Even in an isolated location such as Western Australia, choosing to learn a musical instrument usually begins with selecting one from a range available at your store or school. Choosing to learn the Trautonium begins with research, then deciding how to build or procure one from Germany. One important decision was made early regarding the naming of the instrument to be built, and so in the Australian idiom, it has become known as "The Traut".

6.1. Sounds and Circuits

The instruments built by Trautoniks are considered to be the most faithful reproductions. The low end Volkstrautoniums have a price of 5,499 Euros (around $8,000 AUD), which similarly to the instrument it is reproducing, is some months of the average wage now.

Importantly, the Volkstrautoniums did not include the subharmonic synthesis that Sala developed later, and so many of the later works created for the Mixtur-Trautonium may be unplayable on this model.

Figure 3. A Volkstrautonium, a far simpler instrument than the Mixtur-Trautonium.

Dieter Doepfer is an engineer and designer of modular synthesizers, and was able to work with Oskar Sala prior to his death, in designing modular synthesizer components, which could replicate the functioning of the Mixtur-Trautonium for the replica instrument located in the museum in Vienna. Doepfer are a widely known and used brand of modular synthesizer who have been designing synthesizers since 1977 and pioneered the Eurorack format. Obtaining their modules is comparatively straightforward and affordable.

The three specific modules which they designed to work with others in their range are the A-104 Trautonium Formant Filter, the A-113 Subharmonic Generator and the A-198 Trautonium Manual, and they have suggested how these can be incorporated with their other analogue modules to create a modular Trautonium (Doepfer Musikelektronik GMBH b). The modules range in price, but a complete system can be purchased as modules for around $2,500 AUD.

One notable issue with the Doepfer system is that of a Frequency Shifter. Whilst Doepfer previously made the A-126 Eurorack Frequency Shifter, this has been out of production for some time due to the unavailability of a specific integrated circuit (the CEM3382) which was an essential part of their Frequency Shifter circuit design. This is another example of the variety of issues in the preservation or recreation of old electronic devices.

Whilst both the Trautoniks and Doepfer systems are both attempts to recreate the Trautonium, their differences are multiple. Firstly the Trautoniks instrument is based on the far simpler Volkstrautonium design, including its industrial design which is faithful to the 1930's instrument. The Doepfer instrument has been created as sound synthesizer, not to look historically correct, but to be part of a standard modular instrument, such as is used and toured by musicians around the world even today.

6.2. Robustness and Future Proofing

Whilst an instrument that can sound and adequately be performed as a Trautonium is the principle desired outcome, building an instrument which is able to withstand being transported to performances, and also be connected to other electronic instruments are reasonable features for a musical instrument in the 21st century.

The Trautoniks instrument though having a vintage visual charm, is built into a wooden case that is an intrinsic part of the instrument, and damage to this case is equally damage to the instrument. It is a monolithic system that does not include any MIDI implementation.

Doepfer's modular system can be built into many housings, and they have available modules to connect to other systems via both MIDI and Control Voltage.

These issues along with the critical presence of a Subharmonic Generator make the case for building the instrument from Doepfer modules, with the exception of the Frequency Shifter which is another Eurorack module – Synthesis Technology E560 Deflector Shield.

7. THE TRAUT

7.1. Building

Once the Doepfer modules arrived, a suitable housing was required to build them into. Discussion with another German company Clicks and Clocks followed, and they were able to build a custom set of Eurorack rails which can be installed as one piece into a standard 6RU roadcase.

Doepfer were also able to provide a kit form power supply and buss boards that were installed into the rear of the case. Germany's standard household electrical voltage is very close to Australia's, and all that was needed was an adapter to change from the European plug format to the Australian one.

The assembled instrument does not look anything like the Mixtur-Trautonium, but does bear some resemblance to the controls on the Mixtur-Trautonium nach Oskar Sala, having a shared aluminium faceplate aesthetic. Importantly, this new version is also much smaller than the original and rebuilt Mixtur-Trautoniums, fitting neatly into the 6RU case with the exception of the Manual.

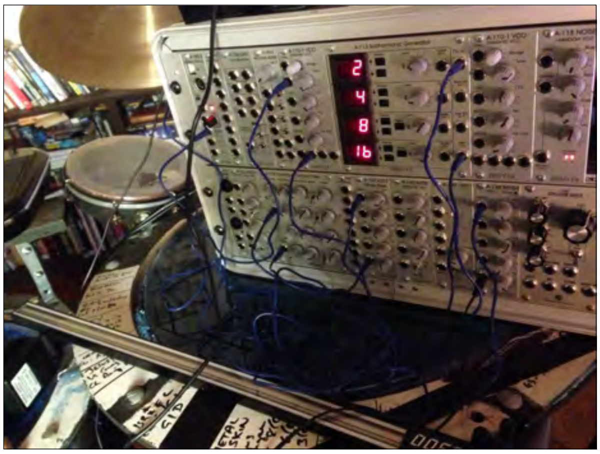


Figure 2. Rebuilt Trautonium from Doepfer modules.

7.2. Archives and learning to play

Re-performing a late 20th century computer music piece for specific electronic instrumentation would require access to readable synthesizer 'patches' as well as the instrument or software to run these on to recreate the correct sounds. Whilst these are known challenges in the practise of digital archiving even of contemporary music, unfortunately this kind of documentation is not generally available for analogue electronic instruments, and even if there were, it is unlikely that files from the 1970's and earlier would be readable now.

Whilst the instrument is now built, as noted Sala did not document what each of the controls on his instrument were, and so a process of listening to the functions of each module on this new instrument is now being followed, along with close listening and viewing of Sala performing, attempting to connect sounds with specific areas of control on the instrument, and translate those into the new system.

Sala left his instrument and seventy shelf meters of archives to the Deutsche Museum, including a large number of tapes of “sound studies” which document his development of many of the classic sounds of the Trautonium, including its most famous performance in Alfred Hitchcock’s *The Birds*.

It is hoped that a combination of research into his notes as well as close listening to recordings of performances and sketches in his archive, will provide a way to recreate the sounds that Sala used in the performance of music for the Trautonium. It is an unusual way to approach learning to play a musical instrument, but though the Trautonium is less than 100 years old, it is the only method available.

8. CONCLUSION

It has been suggested that most electronic and computer based musical instruments are unstable and ephemeral which does not encourage performers to become virtuosos in their playing (Freed, Uitti et al. 2013). Whilst this was intended as a comment on digital instruments, the same is true for analogue electronic instruments.

Tension lies between the need for stable and available instruments for performers to make use of, and the current situation where many of these older electronic instruments are, if accessible at all, on display in museums where the focus is on preserving the machine rather than making it available for use as a performance instrument. For these reasons, re-creating a low cost and robust Trautonium becomes desirable to allow the instrument to be learned by others in the future, and to enable its future development as technology progresses.

9. REFERENCES

120years.net. "The ‘Trautonium’ Dr Freidrich Trautwein. Germany, 1930." Retrieved 7th October, 2015, Available online at URL 120years.net/the-trautoniumdr-freidrich-trautweingermany1930/

Bachimont, B., J.-F. Blanchette, et al. 2003. Preserving Interactive Digital Music: a report on the MUSTICA research initiative.

Deutsche Museum a. "Oskar Sala Fonds - Current Projects." Retrieved 5th November 2015, Available online at URL www.oskar-sala.de/oskar-sala-fonds/oskar-sala-fonds/aktuelle-projekte/.

Deutsche Museum b. "Oskar Sala Fonds - Trautonium." Retrieved 4th November 2015, Available online at URL www.oskar-sala.de/oskar-sala-fonds/trautonium/.

Doepfer Musikelektronik GMBH a. "The Trautonium Project." Retrieved 5th November 2015, Available online at URL www.doepfer.de/traut/traut_e.htm.

Doepfer Musikelektronik GMBH b. "Trautonium System." Retrieved 8th October 2015, Available online at URL www.doepfer.de/a100_man/A100_System_Trautonium_2010.htm

Freed, A., F.-M. Uitti, et al. 2013. “Old” is the new “New”: a Fingerboard Case Study in Recrudescence as a NIME Development Strategy. New Interfaces for Musical Expression.

Giaretta, D. 2011. Advanced Digital Preservation doi:10.1007/978-3-642-16809-3: Springer

International Council of Museums. Museum Definition - ICOM. Retrieved 8th October 2013, Available online at URL icom.museum/the-vision/museum-definition/

Kock, W. 2012. The Creative Engineer, Springer.

Sala, O. 1956. "Synchronizing system for electrical musical instruments." Retrieved 8th October 2015, Available online at URL www.google.com.au/patents/US2740892.

Trautoniks 2012. "Trautoniks References." Retrieved 8th October 2015, Available online at URL www.trautoniks.com/history.html.

Trautonist 2012. Oskar Sala - Frequency Shifter Frequenzumsetzer 1965/1997.

MAKING MUSIC LEARNING FUN: DESIGNING AN INTERACTIVE IBOOK FOR INFORMAL LEARNING

James Humberstone
Sydney Conservatorium of Music
Music Education Unit
University of Sydney

John R. Taylor
Sydney Conservatorium of Music
Composition Unit
University of Sydney

ABSTRACT

The use of new-format tablet computers (such as iPads) in educational settings has increased markedly since the first tablet computer was made commercially available. In secondary schools, the use of tablet computers has been used to help students learn a variety of subjects, particularly numeracy, literacy, and the sciences. The success of using tablet computers in schools is largely owing to the diversity of available electronic resources, such as Apps and web-based content that transform the boundaries between formal and informal learning. This is enhanced by the immediacy of access by students, both at school and at home. However, with the focus of interactive tablet content largely aimed at core subjects, there has been limited interactive content designed solely for music learning. In this paper we describe the design philosophy, and learning and potential teaching benefits of an interactive iBook that has been developed for secondary school students (years 7 and 8) to learn music. A demonstration of the iBook is also provided.

1. INTRODUCTION

Electronic books have been a disruptive force in educational innovation. The benefits of which have been assessed through worldwide pilot studies in a variety of subjects within different levels of education (Couse, Chen and Dora, 2010; Henderson and Yeow, 2012; Jennings et al., 2011; Kiger, Herro and Prunty, 2012). These pilots use different technologies, including the Kindle (Barack, 2011; Martinez-Estrada and Conaway, 2012) and the iPad (Olcese and Murray, 2011; Lau and Siu-lua, 2012). Owing to the design of the devices these technologies offer two different learning experiences: a passive experience similar to a paper book (the standard Kindle text-only experience), and an interactive experience that allows multi-touch, video and audio playback and user input (such as Apple's iBooks). The multifunctional aspect of the iPad, in particular, provides many options for both educators and learners to innovate and enhance their relevant experiences.

Many studies have examined the impact of mobile technology on pedagogy, and have found that not only do teaching practices change with technology (Bebell and Kay, 2010), but a large part of the successful adoption of technology in the classroom lies in the management of the technology (Henderson and Yeow, 2012), not just in terms of the practical management of the devices, but in the integration of the technology with pedagogy and content (Melhuish and Falloon, 2010; Mishra et al., 2011). For students, there is clear evidence that technology use in classrooms increases engagement, interest level, and achievement, as well as enhancing collaboration (Bebell et al., 2014; Bebell and O'Dwyer, 2010; Melhuish and Falloon, 2010). Students were also found to have increased motivation to learn independently (Kinash et al., 2012), which in the case of mobile devices that are accessible outside the classroom, is a critical part of informal learning. This research presents a set of clear design parameters when developing an interactive tool for learning music. This is shown in Figure 1.

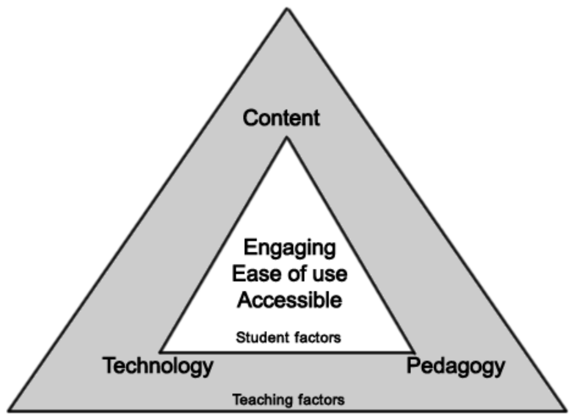


Figure 1. iBook design factors.

From a teaching perspective, the design of the iBook must be relevant in terms of content, technology, and pedagogy, in which these factors relate more to learning in a school environment. These factors should be considered in terms of appropriateness for the age group, curriculum, pedagogy and teaching philosophy, technological constraints and institutional policies that may promote or inhibit the use of such technology. Although teaching factors are a crucial component in the technology integration mix, there are student factors that relate to the success of the technology in both school, and out-of-school (informal) learning environments, with informal learning (in this case independent learning) being a critical component of a student's learning experience. A student's potential willingness or motivation to learn away from the classroom environment can be enhanced by addressing the student factors and appealing to the student: by ensuring that the iBook and the content is engaging, easy to use, and easy to access outside of class.

This paper will focus on describing the design of an interactive music learning resource as an informal learning tool, from a student-centred design perspective. Although this approach omits teaching level factors, the synergy between student engagement and independent learning allows greater freedom to explore the interaction design opportunities on the iBook platform, for improving teaching and learning.

2. ADOPTED PEDAGOGY

Given the repeated themes of student motivation, independent learning and autonomy outlined in the literature around mobile learning, it was decided to adopt an established music pedagogy that also promoted these facets of learning.

Lucy Green's seminal text *How Popular Musicians Learn* (2002) included a series of case studies that explored that very topic, and summarised key learning features such as learning aurally by copying rather than from notation, learning in friendship groups, and learning music of their own choice. It established the term *informal learning* in music education.

In *Music, Informal Learning, and the School: a New Classroom Pedagogy*, Green (2008) extended this research with further case studies which transformed the initial findings into a pedagogical structure that could be used in schools. It was so successful that it was adopted as the model for the British intervention on music education called *Musical Futures* (Ainscough et al., 2009) which further consolidated Green's extensive research to a number of formulas that music educators could follow to bring informal learning through their music curriculum. Musical Futures was incredibly successful, with several studies showing that it improved student engagement, enhanced learning and enjoyment, and resulted in two to three times the number of students choosing to study music further (Hallam et al., 2009, 2011).

2.1. Modelling Aural Learning

The approach "Modelling Aural Learning" was adopted from Green (2008) and Ainscough et al.'s (2009) texts. In this variation on the informal learning pedagogy, students do not choose their own music, but the teacher selects music which is culturally relevant to students and which in their professional judgement they believe brings students great opportunity for musical learning.

Teachers are required to provide audio tracks of individual parts in the selected song, so that students do not have to pick one part out of a mastered song. Notation can also be provided, but it must be suitable for the genre and instrumentation provided: for example, guitar chord diagrams or tablature rather than traditional music notation.

Throughout development of the iBook content, the Modelling Aural Learning approach was followed.

3. DESIGN CONCEPTS

3.1. Interactive Elements

Apple's iBooks Author allows an iBook to exploit the multi-touch functionality of the iPad using widgets, to control a range of multimedia elements that can be embedded into the iBook. These multimedia elements include photo galleries, chapter reviews, audio, video, 3D objects, interactive images, "pop-overs", links and anchors, and html. Because "widgets" can be custom programmed using HTML, CSS and JavaScript, functionality can be extended. Such interactive elements have been utilised to great effect in educational iBooks such as E. O. Wilson's *Life on Earth* series (Wilson et al., 2014). In the *Life on Earth* series, the interactive elements assist the reader in understanding more complex scientific concepts, and are designed as sub-stories/explanations within the main theme of the book.

In contrast, using interactive elements to teach music performance and composition requires a different approach, revolving around the use of exemplars. For example, music performance can be taught using video performances of small excerpts, whereas video examples of some scientific concepts may prove too difficult. Similarly, an animation may be suitable to demonstrate cell division, but would be impractical to demonstrate fingering of long complex musical phrase on an instrument. In such a scenario, it is important to identify which interactive elements convey a concept in the most engaging and understandable way. Furthermore, the interactivity should be used in such a way as to be complimentary to the structure of the iBook.

3.2. Structure of the Book

The aim of this book is to teach students the fundamental elements of music composition and to develop their music performance skills by presenting new material for them to learn. In a sense, the iBook is intended to introduce a new structure for music learning, or a new context in which to apply skills previously learned. One such method for structuring such learning is by using "instructional scaffolding" (Bull et al., 1999). Bull et al. describe scaffolding in the context of technology in learning environments, and present concepts of scaffolding that can help not only organize the general structure of the book, but to organize the types of interactive elements that can be used to augment the various stages of learning. Bull et al describe these concepts as: advance organisers (the link between what is known and the new knowledge), modelling the appropriate behaviour), exploring new skills, and being generative with the learned information.

3.2.1. Advance organisers

One could argue that the iBook itself is an advance organizer, since it presents new information to learners. Conceptually, a positive way to motivate and engage students is to construct a scenario that can easily be related to. In the case of music, this may be situating the book in context of an individual artist or genre of music. This information could be presented using audio, video and photos in order to make the 'iBook story' or theme more realistic.

3.2.2. Modelling

In real-world music learning, particularly instrumental learning, students often take part in one-to-one lessons with the teacher. In these lessons, the teachers will demonstrate the tasks (and/or behaviours) in a step-by-step manner that the students are expected to learn. Since the iBook will be a substitute teacher (during periods of informal, independent learning), it is important that these tasks continue to be demonstrated. One way that this can be done is through the use of video examples, either embedded in the iBook, or as hyperlinks to online video sites.

3.2.3. Exploring new skills

Teachers often encourage students to explore tasks with a view to the students accomplishing that task. In a musical context, one example might be for a student to explore the use of a scale or the exploration of a chord sequence. Exploring new skills using an iBook could include setting small challenges using 3D objects, animations, pop-overs, or even links to external online tests. It could also include getting students to play along with a video or audio excerpt, or by encouraging students to make recordings of their performances using third party apps.

3.2.4. Being generative with learned information

Students become generative with learned information when they finally have control. In the case of music, this could relate to learning a new chord sequence, and demonstrating to another learning, or by generating new contexts for the learned skill. This particular aspect is closely related to composition, since it aligns closely with understanding musical concepts and generating new music from those constructs (Hickey, 2003; Kaschub & Smith, 2013). In the context of an iBook, this could include an interactive graphic that asks students to re-order musical fragments, or to play a scale to music of a different genre. It could also include setting tasks that involve using third party apps to record or compose music. This work than then be presented in the school environment or collaborated on with peers.

4. DESIGN THEME: ELISHA KEEN

4.1. Developing a musical concept

The musical theme of the iBook is of paramount importance for student engagement. With significant number of choices in musical genre and musical context available to choose from, it is important to find a balance between student engagement, musical complexity and accessibility. The Modelling Aural Learning approach also dictates that music selected must be relevant to students' own musical cultures. Choosing a narrow theme with limited appeal, e.g. an unpopular genre of music, has the potential to limit student engagement with the learning resource (Hein, 2013). In addition, a genre of music that has a higher perceived complexity may also have similar detrimental effects. Both of these aspects of informal learning are related to the conceptual accessibility of the resource in terms of the link between what is known, and familiar, with the new knowledge. In other words, students should be at least familiar with the genre. The accessibility of the resource for the student extends beyond knowledge linking and advance organization, into a priori cultural, personal and individual experience, and includes gender neutrality, and aspirational motivation. Consequently, the music of singer/songwriter/performer Elisha Keen was chosen as a case study for the iBook. In order minimize the complexity of the iBook, the resource examines her song "Breathe Some Different Air". This song is a contemporary ballad that uses a typical popular song structure (verse/chorus/verse etc.).

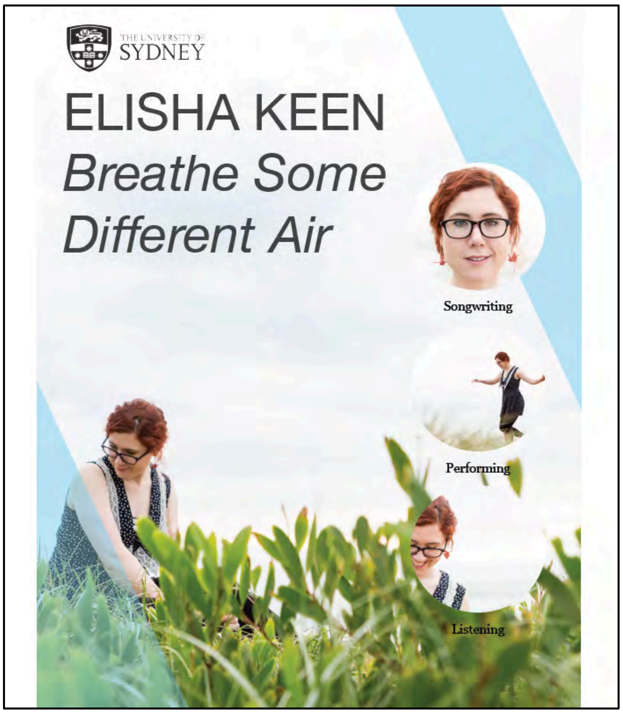


Figure 2. The front page of the iBook.

To increase Elisha Keen's accessibility with students, a video of Elisha talking about her songwriting processes are embedded into the early pages of the iBook. This has a dual purpose. First, Elisha discussing her songwriting processes is presenting new information to students. It seeks to motivate and encourage students to continue. Second, the video provides a modelling exemplar, whereby the process of composing a song is demonstrated to students. Furthermore, the interview was filmed in the recording studio to add aspirational qualities to the narrative.

Figure 3. An iBook page with an embedded video interview with Elisha Keen.

4.2. Modelling and Exploring using video and audio

In Modelling Aural Learning, in the absence of a teacher providing examples in person and teaching didactically, it is still possible to provide a mechanism for demonstrating, or modelling, the tasks or skills. Since users of the resource may have different skill levels, and may not be familiar with the song to be played, it is important that the modelling continues the step-by-step demonstrations. In this iBook, a layered approach was taken, where modelling was designed to incrementally add examples that relate to different levels of compositional complexity. The interactive element chosen for the modelling of the first layer, is video of how to play the chord progressions on the guitar. This video is accompanied by tablature outline the chord progressions. One of the benefits of using this resource lies in the ability for the learner to play back the video until they have learnt the sequence.

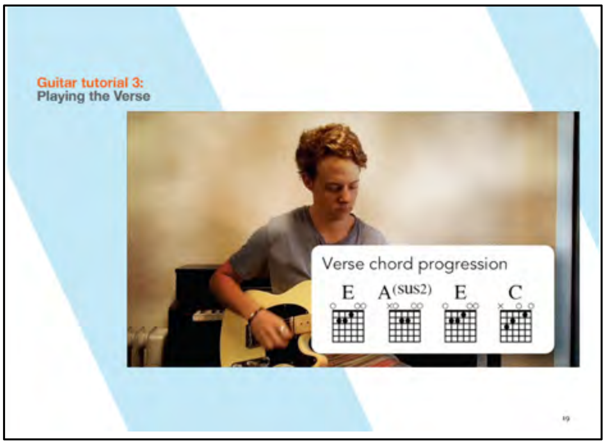


Figure 4. Modelling of the instrumental performance is done using video and graphics for the individual sections of the song.

In order to assist in the modelling process, students are given exemplars in the form of audio files of individual parts that demonstrate the different sequences. Elisha Keen provided the original stem recordings, so that examples are the same high standard as the final mixed song. This gives students the opportunity to explore their knowledge of the individual chords in the context of longer sequences of chords and with different instrumental contexts, or musical layers. This modelling approach is continued and expanded upon to include different instruments, including the vocal tracks, bass guitar and drums, with minimal notation.

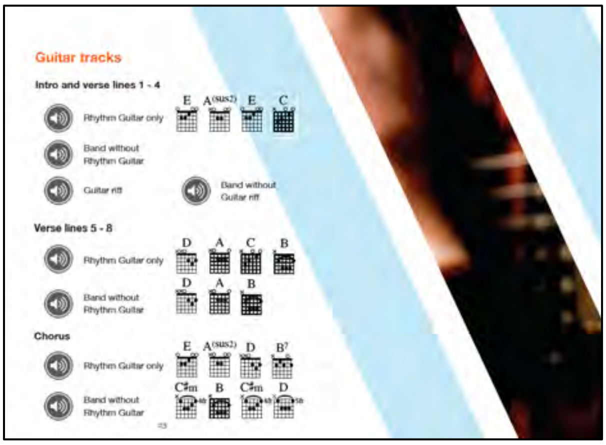


Figure 5. Students are given exemplars of the guitar sections in the form of audio examples, and guitar tablature for them to play along.

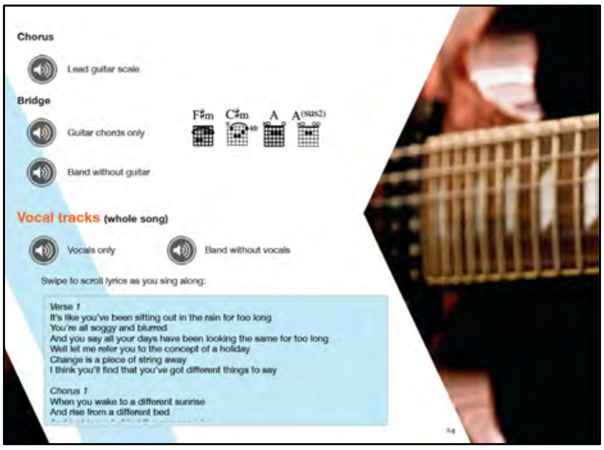


Figure 6. Students are also given vocal track exemplars.

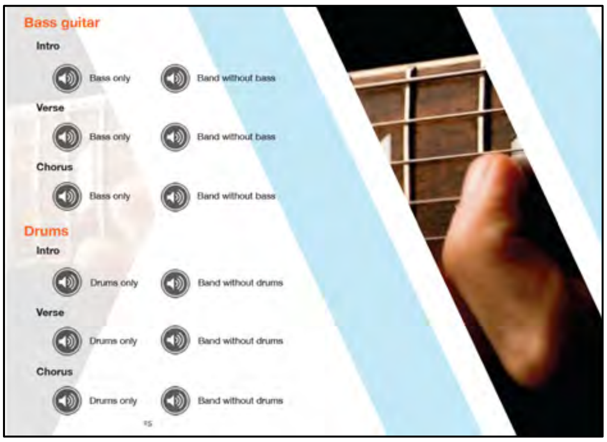



Figure 7. Students are given bass guitar and drum exemplars.

By adding the exemplars of additional instruments, students are given the opportunity to hear different elements of the song in isolation. It also increases the accessibility of the song as they are exposed to the constituent components without interference, in order to assist in the development of additional skills, such as listening. Furthermore, this format encourages the student to explore newly learned skills by giving students the opportunity to play along with individual instruments, even though this may not be expressly stated in the iBook. Once the different sequences of chords have been learned, these sequences are contextualized in chord sheet that matches the chord sequences with the lyrics. This adds a final layer of complexity to the learning of this song.

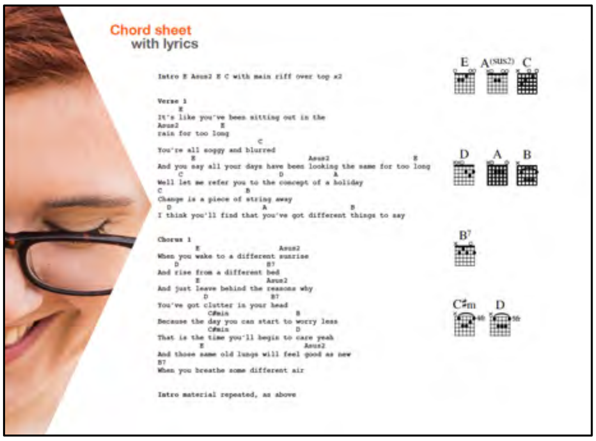


Figure 8. Students are given a chord sheet that relates the chord sequences to the vocals.

4.3. Being generative using interactive animation

The final part of the iBook enables students to be generative with the learned information, specifically in relation to the song structure. In order to allow students to expand create new sequences, an interactive animation allows the student to select the different parts of the song, reorder them, and listen back to it.

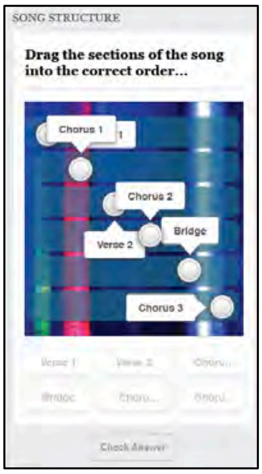


Figure 9. A task for students to explore new information. Students are asked to re-order sections of the music to see what the effects are.

This activity was designed as an exploratory tool to enable students to gain an understanding of song structure, and to create and listen to new structures. Furthermore, this activity tests the students' knowledge and engages students in thinking about the song from a structural and holistic perspective.

5. CONCLUSIONS

As a starting point, the iBook content being informed by key aspects of informal learning pedagogies creates synergies with the framework for instructional scaffolding, whereby many factors that that are designed to engage, encourage and increase accessibility of the content are inherent in the interactive objects. For example, the video tutorials with graphics on how to play a chord sequence, and the interactive animations for re-ordering song sections both inherently engage, encourage and make the content accessible in a manner that is familiar to music learning.

By drawing upon concepts of instructional scaffolding, the iBook clearly presents a narrative. Content that serves as advance organisers are augmented by interactive elements to ensure clarity, and relevance to informal and independent learning. The exemplars that demonstrate how to play various parts of the song not only model appropriate behaviour and skill, but as a collection of exemplars across different instruments allow greater levels of conceptual complexity, while simultaneously allowing students the opportunity to explore new skills in different contexts.

The separation of musical elements from the song also allows students with different instruments to use the resource, and to enhance their learning. Furthermore, this approach enables students learning different instruments to collaborate musically by either performing together, composing together, or simply by teaching other students. Finally, the interactive animation is the most appropriate method for encouraging students to be generative with learned information.

One conclusion that can be drawn from the iBook design framework, is that there are at the very least some basic principles in interactive iBook design that should always be considered. This relates specifically to the type of interactive element used to convey a particular teaching element. In this paper, videos have been used in predominantly the iBook narrative and instructional modelling. In these two areas of the iBook, students are "learning". In other areas, such as exploring new skills and being generative with learned information, students are "participating" (even though they are still learning). In contrast, audio is predominant throughout the iBook, and as a ubiquitous element, transcends instructional boundaries. Thus, it is worth considering the objective outcome of each section within the context of the learning experience, as to the type of interactive object used.

6. FURTHER WORK

Further work is underway to develop the iBook resource to include composition frameworks, testing, and teaching factors, particularly in relation to curricular alignment, in order to evaluate the resource as a formal learning tool. In addition, further work could include trialling the iBook in schools to assess the impact of such electronic resources on teachers and students in music education. This could include extending the iBook content to different themes, including different genres of music, or including music from different cultures.

7. REFERENCES

Barack, L. 2011. “The Kindles Are Coming: Ereaders and Tablets Are Springing up in Schools - And Librarians Are Leading the Way.” *School Library Journal,* 57(3): 58-60.

Bebell, D., & Kay, R. 2010. “One to One Computing: A Summary of the Quantitative Results from the Berkshire Wireless Learning Initiative.” *Journal of Technology, Learning, and Assessment, 9*(2). Retrieved February 17, 2013, from http://www.jtla.org.

Ainscough, M., Burton, I., Crawford, H., Cross, J., D'Amore, A., Davis, C., and Wolfe, A. 2009. *Musical Futures. An approach to teaching and learning*, ed. Abigail D'Amore. London: Paul Hamlyn Foundation.

Bebell, D., & O'Dwyer, L. 2010. “Educational Outcomes and Research from 1:1 Computing Settings.” *The Journal of Technology, Learning and Assessment,* 9(1): 16. Retrieved February 17, 2013, from http://www.jtla.org.

Bebell, D., Clarkson, A., & Burraston, J. 2014. “Cloud computing: Short term impacts of 1:1 computing in the sixth grade.” *Journal of Information Technology Education: Innovations in Practice,* 13, pp. 129-151.

Bull, K. S., Shuler, P., Overton, R., Kimball, S., Boykin, C., and Griffin, J. 1999. “Processes for developing scaffolding in a computer mediated learning environment.” In *Rural Special Education for the New Millennium. Conference Proceedings of the American Council on Rural Special Education (ACRES),* Albuquerque, New Mexico, pp. 241-248.

Couse, Leslie J., and Dora W. Chen. 2010. “A tablet computer for young children? Exploring its viability for early childhood education.” *Journal of Research on Technology in Education,* 43 (1): 75.

Green, L. 2002. *How Popular Musicians Learn: A Way Ahead for Music Education* (Aldershot: Ashgate Publishing Limited 2002), Ebook PDF edition.

Green, L. 2008. *Music, Informal Learning and the School: A New Classroom Pedagogy*. Aldershot: Ashgate Publishing Limited.

Hallam, S., Creech, A., Sandford, C., Rinta, T., and Shave, K. 2009. *Survey of Musical Futures*: *A Report from the Institute of Education,* London: University of London. Retrieved from www.musicalfutures.org: https://www.musicalfutures.org/resource/27229/title/instituteofeducationreportintothetakeupandimpactofmusicalfutures.

Hallam, S., Creech, A., and McQueen, H. 2011. *Musical Futures: A case study investigation, Final Report,* London: University of London. Retrieved from: https://www.musicalfutures.org/resource/27646/title/instituteofeducationlongitudinalstudyofmusicalfutures.

Hein, E. “Designing the Drum Loop. A constructivist iOS rhythm tutorial system for beginners” (Master's diss., New York University, 2013).

Henderson S., and Yeow, J. “iPad in Education: A Case Study of iPad Adoption and Use in a Primary School.” *Proceedings of the IEEE 45th Hawaii International Conference on System Sciences.* Hawaii: IEEE, pp. 78-87.

Hickey, M. (ed.). 2003. *Why and How to Teach Music Composition: A New Horizon for Music Education*. Reston: MENC: The National Association for Music Education.

Jennings, G., Anderson, T., Dorset, M., and Mitchell, J. 2011. *Report on the Step Forward iPad Pilot Project,* Melbourne: The University of Melbourne. Retrieved January 4, 2011, from http://www.trinity.unimelb.edu.au/Media/docs/iPadPilotReport2011-1b1e1a52-79af-4c76-b5b6-e45f92f2c9e9-0.pdf.

Kaschub, M., and Smith, J. 2013. *Composing our Future: Preparing music educators to teach composition*. New York: Oxford University Press.

Kiger, D., Herro, D., and Prunty, D. 2012. “Examining the influence of a mobile learning intervention on third grade math achievement.” *Journal of Research on Technology in Education,* 45(1): 61.

Lau, APT., and Siu-lau, H. 2012. “Using iPad 2 with note-taking apps to enhance traditional blackboard-style pedagogy for mathematics-heavy subjects: A case study.” In *Proceedings of IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)* 2012. H3C-4-H3C-6.

Martinez-Estrada, P. D., and Conaway, R. N. 2012. “EBooks: The Next Step in Educational Innovation.” *Business Communication Quarterly,* 75(2): 125-135.

Melhuish, K. & Falloon, G. (2010). “Looking to the future: M-learning with the iPad.” *Computers in New Zealand Schools: Learning, Leading, Technology*, *22*(3): 1-16.

Mishra, P., Koehler, M.J., and Henriksen, D. 2011. “The seven trans-disciplinary habits of mind: Extending the tpack framework towards 21st century learning.” *Educational Technology, 11*(2): 22-28.

Murray, O., and Olcese. N. 2011. “Teaching and learning with iPads, ready or not?” *Techtrends.* Vol. 55. Boston: Springer US.

Wilson, E., Morgan, R., and McGill, G. 2014. *Life on Earth (Units 1-7).* (Cambridge, MA: Wilson Digital Inc. 2014). iBook edition.

CARME: A MULTI-TOUCH CONTROLLER FOR REAL-TIME SPATIALISATION OF GESTURAL BEHAVIOUR

Blake Johnston
Victoria University
Wellington, New Zealand

Michael Norris
Victoria University
Wellington, New Zealand

Bridget Johnson
Victoria University
Wellington, New Zealand

Ajay Kapur
Victoria University
Wellington, New Zealand

ABSTRACT

Expressivity has been the primary driver of recent research and development in the field of mechatronic musical systems. In the search for greater expressivity, however, the complexity of systems tends to increase concomitantly, thanks to the proliferation of control parameters, and the desire to group multiple mechatronic units into 'ensembles'. This then gives rise to the problem of how to control such a complex system in a real-time and intuitive manner. This paper proposes some novel solutions to the conceptual and logistical problems inherent in the 'data explosion' of expressive mechatronic systems, through the context of the design and development of a multi-touch app, *Carme*, custom-built to control *The Polus Ensemble,* a mechatronic sound sculpture ensemble. The potential for *Carme* to control other ensembles as well as sound-spatialisation systems is also considered.

1. INTRODUCTION

1.1. The Polus Ensemble

The Polus Ensemble is a collection of six mechatronic sound sculptures, each with a single string and a bowing mechanism (Johnston et. al. 2014). Its design focuses on two main considerations: a) allowing for a range of musical controls, and b) that the objects should be both physically and visually compelling, whilst blurring and relating the sonic, visual and physical characteristics. For more detailed explanation of the aesthetics and motivations see Johnston et. al. 2014.

Jim Murphy defines expressivity in the context of mechatronic instruments as 'the ability of a mechatronic musical system to affect a wide range of musical parameters' (Murphy 2013). The desire for each sculpture to be musically expressive led to designing a system that afforded control of: a) the pressure of the bow on to the string, b) the speed and direction of the bow, and c) whether or not the string was dampened.

The Polus Ensemble draws upon both the mechatronic instrument field, as well as the mechatronic sculpture field, blurring the line between the two. The range of controllable parameters in each unit plays a part in the musical expressivity of the ensemble, although the ensemble as a whole should be seen as one mechatronic system. Therefore, part of the complexity comes from the spatial distribution and relationship between each unit.

Although *Carme* was developed to specifically accommodate the controllable parameters available in *The Polus Ensemble*, it has been designed in such a way that it could also control other, similar, mechatronic systems.

Figure 1. The largest of *The Polus Ensemble* sculptures. Provided by Victoria University.

1.2. Mechatronic Instruments

The field of mechatronic instruments has blossomed in the recent past, as the technology to rapidly prototype and create inexpensive systems has become more readily available. These new instruments tend to draw upon a conventional instrumental design paradigm, being influenced by, or even emulating, existing instrument designs.

A comprehensive overview of the field is outside the scope of this paper,[1] although two exemplars of complex, expressive mechatronic instruments can be seen in Jim Murphy's *Swivel 2* (Murphy 2013) and *MechBass* (McVay et. al. 2011), the latter collaboratively built with James McVay. *Swivel 2* is a six-string mechatronic slide guitar with a rotating arm, which allows for different pitches to be stopped, and the pressure against the string varied. Furthermore, each string can be plucked by a rotating plectrum, with a range of force possible. *MechBass* is a four-stringed mechatronic bass guitar, affording a fine degree of pitch control, picking velocity and dampening. Instead of the rotating arm of *Swivel 2*, linear fretting mechanisms move along each string to control pitch. A similar rotating plectrum design is used.

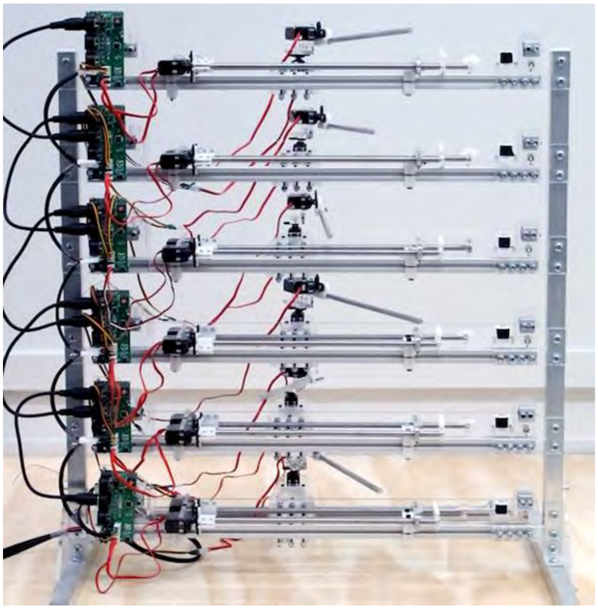


Figure 2. Murphy's mechatronic slide guitar, *Swivel 2*.

Both *MechBass* and *Swivel 2* have a wide range of control parameters to afford greater musical expressivity.

1.3. Mechatronic Sound Installations

As well as conventional instrumental paradigms, there are many examples of mechantronic technology being used in sound installation art. From simplistic units to complex kinetic sound sculptures, the integration of mechatronic systems is widespread in sound-based art. A full overview of the field of mechatronic sound installations is outside of the scope of this paper, although the work of Swiss artist Zimoun is illustrative of similar complex control problems that emerge from this field.[2]

Zimoun creates complex emergent systems from large arrays of simplistic mechatronic elements. DC motors with attachments are often used to strike a physical object, as in *329 prepared dc-motors, cotton balls, toluene tank* and *157 prepared dc-motors, cotton balls, cardboard boxes 60x20x20cm* (Zimoun 2013, 2014).

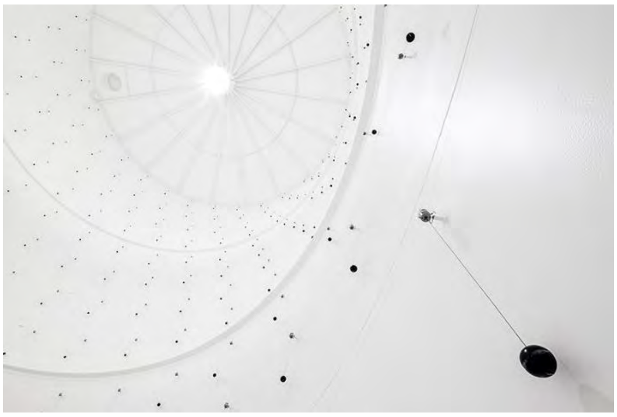


Figure 3. Zimoun: *329 prepared dc-motors, cotton balls, toluene tank*

Martin Messier combines existing technology with mechatronic control systems in his works *Projectors* and *Sewing Machine Orchestra.* In a similar vein, Nicolas Bernier has used tuning forks with mechatronic actuating systems in *frequencies (a)* (Messier 2012, 2014; Bernier 2012).

Figure 4. Martin Messier – *Sewing Machine Orchestra*

Mo H. Zareei has produced a series of mechatronic noise-intoners, the *Brutalist Noise Ensemble*, consisting of three modules: *Rippler, Rasper,* and *Mutor* (Zareei 2015). Each module is a series of noise-intoners that use

[1] For a comprehensive overview of the mechatronic instrument field: Kapur 2005; Murphy 2012, 2015; Maes 2011; and Singer et al. 2004.

[2] For overviews of sound art that include mechatronics: Licht 2007, LaBelle 2006, Kim-Cohen 2009 and Voegelin 2010 offer examples. At time of writing, there does not seem to exist a comprehensive overview of mechatronics in sound art.

simplistic mechatronic parts to create highly controllable bursts of noise. With a focus on rhythmic structuring of noise, the ensemble as a whole is capable of producing complex, rhythmic noise works.

Figure 5. Mo H. Zareei – *Brutalist Noise Ensemble*

1.4. Control Systems for Mechatronics

In comparison to the developments in these mechatronic systems, the state of the controller field has been far more stagnant. As Murphy recognizes,

> ‘although expressive control allows for a wide range of musical possibilities for the composer to explore, a pressing need for increased user friendliness in human-robot interactions has been identified: To compose for these systems in their current state is to manually direct every action that the robot undertakes. To write music in this manner is quite time-consuming, requiring much actuator management rather than higher-level musical composition. (Murphy, 2015)’

1.4.1. Existing Controllers for Live Performance

Most of the mechatronic systems use MIDI as the basis of their communication design. While this means that many commercial MIDI controllers are compatible with their communication system, allowing for off-the-shelf real-time control, MIDI controllers are not designed to specifically control mechatronic systems and, thus, are not well suited to interacting with the more complex, continuous and non-linear mechatronic systems.

1.4.2. Non Real-time Composition

Due to the limitations of existing controllers, composers often opt for a software solution and compose MIDI scores to be played back. While this allows a high degree of control, and ensures repeatability through playback, it imposes many limitations on the interaction a user can have with the system. For instance, only expert users who have learned the specifics of how each control parameter is mapped will be able to control the mechatronic system with any degree of precision. This can be a cumbersome method, even for these expert users, as they try to deal with a large number of parameters at once. This method also limits real-time interaction, which makes it ill-suited for real-time performance, or a public, interactive installation setting, in which untrained, non-specialist audience members may be invited to provide some degree of interactive control.

This lack of intuitive, real-time control system was a strong driving force for designing *Carme* for the *Polus Ensemble*, in an attempt to realise the interesting interactions possible with a mechatronic system in both a performance and installation setting.

2. SYSTEM OVERVIEW

2.1. Designing for The Polus Ensemble

Each sculpture in *The Polus Ensemble* has a bowing mechanism that allows for a range of control. This control is manifested through four controllable parameters: arm direction, arm speed, bowing wheel direction, and bowing arm speed.

The bowing wheel acts as a continuous bow, rotating as it is brought into contact with the string by the arm. Once in contact with the string, the arm can also control the amount of pressure of the wheel against it. This pressure, in combination with the speed of the bowing wheel, controls the loudness of the sculpture—adding depth to the possible sounds that can be created.

The design of the bowing mechanism allows for a small range of behaviours, each with some expressive control. A basic continuous tone can be controlled, and the speed of the wheel vs the pressure against the string can be explored—resulting in a continuous range of timbres between overpressure (‘scratch tone’) and light pressure (*flautando*). Through changing the direction of the bowing wheel during a held note, rhythms can be articulated and, with fast repetitive changes, a tremolo effect can be achieved. A percussive strike can also be performed by controlling how the bow approaches the string, quickly striking and sounding.

With six sculptures in *The Polus Ensemble*, this design allows for 24 control parameters in total, a large number to attempt to control in real-time.

Figure 6. *The Polus Ensemble*

2.2. Communication Framework

Each sculpture has its own set of electronics attached to the instrument. This includes an Arduino, motor driver shield, and a custom-made PCB board that allows for MIDI communication, and a data feedback loop for motor control.

The ensemble is chained together through MIDI INs and MIDI THRUs, with each unit having a separate channel to route messages. The firmware on each electronics module is directly passing the MIDI information on to the four control parameters, allowing for more complex and configurable control to be done on the computer

The MIDI information received by each electronics module is sent from a computer that is running Max/MSP. This mid-point is where the majority of the data handling and processing occurs.

The Max/MSP application receives OSC data, wirelessly sent from the iPad. OSC was chosen as its highly configurable, and with the possibility of transmitting data wirelessly, the control chain from the computer to the ensemble can be isolated, allowing for the user to focus on the interaction with the ensemble.

Figure 7. Control flow of *Carme* communication framework

2.3. Hardware Platform

The iPad was chosen as the hardware platform for three main reasons. Firstly, the iPad, and other touch based devices, have become widespread and common. This means that for most people, interacting with an iPad is not confronting, and there is an understanding of possible interactions with it. This is essential in an interactive gallery setting to encourage audiences to interact with a work.

Secondly, the popularity of the device means that the application can be potentially used with other mechatronic systems easily, without the need of building a new physical interface.

Finally, the iPad allows for a very reliable, accurate multi-touch hardware platform, with built-in gestures that can be used to add functionality.

3. APP DESIGN

3.1. Design Approaches

The app is designed with a simple 'spatialized behaviours' model, which, while to some degree abandons the possibility of fine-grained control over every possible control parameter, has the benefit that it greatly simplifies the user experience, while still allowing for complex behaviours to result. The design approach for *Carme* considered the two different complexities that exist in *The Polus Ensemble*: 1) the spatial distribution of multiple sculptures, treated as one mechatronic system; and the multiple parameters of control that each sculpture presents.

The app expands on the spatial concepts of *The Polus Ensemble* by using the virtual space of the iPad, and mapping this with the physical arrangement of *The Polus Ensemble*. Spatial relationships in the virtual space of the iPad are realised by creating sound, actuated in the real, physical space of the distributed sculptures.

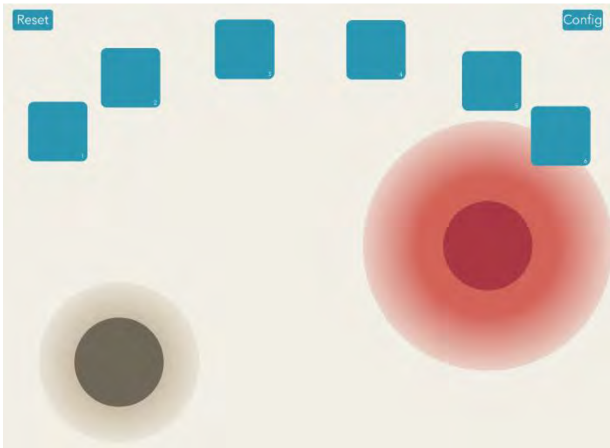


Figure 8. Carme showing six instruments for the Polus Ensembles and two behaviours: one expanded and overlapping with an instrument, and another to the side.

Two major types of abstract shapes are shown. The rectangles represent the individual instruments, while the circles are behaviours.

Due to each unit having four raw control parameters, behaviours were abstracted to allow a more intuitive interaction. This means if a user wanted to create a held continuous note, they would not have to set the direction of the arm and adjust the speed, as well as, control the speed of the bowing wheel. Instead these behaviours offer an abstract control like intensity, which are interpreted in the Max/MSP application to control these parameters accordingly.

By dragging a behaviour in close proximity to an instrument, the relevant behaviour is created in the instrument. This design can easily be customized to have any amount of instruments and behaviours on the screen, catering to the specific mechatronic system. This behaviour is only triggered when the shapes overlap. The degree of proximity is used to control a parameter within the behaviour. An example would be the intensity of a continuous note.

Each shape is not fixed in space and may be grouped in any way. This allows for an interesting approach to organising sound, as sculptures can be spatially grouped together to create chords, or separated to make small sections in the ensemble.

The size of the behaviours can be altered by using a spreading gesture with two fingers. This allows for the influence of a behaviour to dilate and cover a large space. By making the behaviours larger, the user can have finer control over the proximity, as the resolution is effectively higher.

A configure screen is also implemented to set the wireless communication parameters of the specific network. Two parameters need to be set accordingly, the I.P. address of the target device, and the port that OSC will be sent over.

4. EVALUATION AND LIMITATIONS

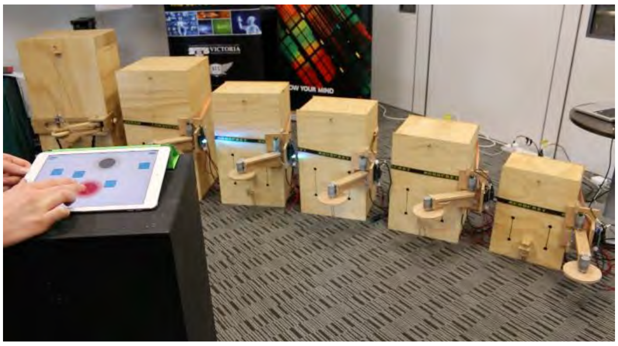


Figure 9. *Carme* with the *Polus Ensemble* in a public setting.

So far, the application has been used in a public demonstration at Victoria University. The application allowed for two behaviours to be spatialized through the ensemble: a held continuous note, and a percussive strike.

Through using the application, and designing behaviours to be controlled, the strengths and weaknesses were revealed in this approach to controlling a complex mechatronic system. For instance, the spatialisation of a simple behaviour around the ensemble worked well. The bowing behaviour allowing A for interesting chords being created through grouping the instrument shapes together and overlapping the bowing behaviour. The percussive strike behaviour created interesting rhythmic interplay through the separation and spatial distribution of short pitched hits.

However, more complex behaviours were less intuitive and not as well suited to this method of interaction. A tremolo bowing behaviour, for instance, could possibly have a parameter for speed of tremolo, as well as loudness, but due to the simplicity of the spatial interaction between an instrument shape and behaviour shape, only one continuous data stream is given per interaction. This means that only one of these parameters can be controlled by a user, and makes trying to control other similar parameters problematic.

A complex system like Murphy's *Swivel 2* that has continuous control over pitch as well as picking intensity would be problematic for this spatial approach. The application could be mapped so that instead of representing a physical space, an abstract pitch space could be explore, although it may be that another approach is needed to intuitively control those systems.

5. FUTURE WORKS AND APPLICATIONS

In the evaluation section, the strengths of this application were explicated, allowing for the potential applications of *Carme* controlling mechatronic systems beyond *The Polus Ensemble* to be imagined. What *Carme* does well is allow for a non-linear spatialisation of simplistic gestures through a possibly large array of mechatronic units.

Works like those of Zimoun—complex through the amount of simplistic modules distributed through space—are strongly suited to being controlled by this approach. This interaction afforded by the spatial design approach would allow an audience to interact in a way that would be heretofore not possible. This spatial complexity can be seen in many other mechatronic works including that of Zareei, Bernier and Messier.

A range of settings are planned that explore the spatial relationships of *The Polus Ensemble*. This includes a performance, controlled by *Carme*, with the ensemble stacked vertically, creating a near four meter tall structure; exploring the vertical spatialisation of pitch space in the overtone series.

What *Carme* affords is not only a way of performing with large, complex spatial mechatronic systems; but a tactile way of giving the audience the power of interaction, in a way that is intuitive and approachable.

6. REFERENCES

Bernier, Nicolas. 2015. "Frequencies (a)." Available Online at www.nicolasbernier.com/page/works.htm. Accessed September 2015.

Johnston, Blake, Henry Dengate Thrush, Ajay Kapur, Jim Murphy, and Tane Moleta. 2014. "Polus: The Design and Development of a New, Mechanically Bowed String Instrument Ensemble." In *Proceedings of the International Conference on New Interfaces for Musical Expression.* Goldsmiths, University of London.

Kim-Cohen, Seth. 2009. *In the Blink of an Ear: Toward a Non-Cochlear Sonic Art*. 1 edition. New York: Bloomsbury Academic.

LaBelle, Brandon. 2006. *Background Noise: Perspectives on Sound Art.* A&C Black. London: Bloomsbury Publishing.

Licht, Alan. 2007. *Sound Art : Beyond Music, between Categories*. New York, NY: Rizzoli International Publications.

Maes, Laura, Godfried-Willem Raes, and Troy Rogers. 2011. "The man and machine robot orchestra at logos." *Computer Music Journal* 35(3): 28-48.

McVay, James, D. A. Carnegie, J. W. Murphy, and A. Kapur. 2011. "Mechbass: A Systems Overview of a New Four-Stringed Robotic Bass Guitar." *Proceedings of the*

19th Electronics New Zealand Conference (ENZCon), Dunedin, New Zealand, pp. 145–50.

Messier, Martin. 2011. “Sewing Machine Orchestra” Online at www.mmessier.com/projects.html. Accessed September 2015.

Messier, Martin. 2014. “Projectors” Online at www.mmessier.com/projects.html. Accessed September 2015.

Murphy, Jim, Ajay Kapur, and Dale Carnegie. 2012. “Musical Robotics in a Loudspeaker World: Developments in Alternative Approaches to Localization and Spatialization.” *Leonardo Music Journal* 22: 41–48.

Murphy, Jim. “Expressive Musical Robots: Building, Evaluating, and Interfacing with an Ensemble of Mechatronic Instruments,” PhD diss. Victoria University, Wellington, New Zealand, 2014.

Murphy, Jim, James McVay, Paul Mathews, Dale A. Carnegie, and Ajay Kapur. 2015. “Expressive Robotic Guitars: Developments in Musical Robotics for Chordophones.” *Computer Music Journal* 39(1): 59–73.

Singer, Eric, Jeff Feddersen, Chad Redmon, and Bil Bowen. 2004. “LEMUR’s Musical Robots.” In *Proceedings of the 2004 Conference on New Interfaces for Musical Expression*, 181–84.

Voegelin, Salome. 2010. *Listening to Noise and Silence: Towards a Philosophy of Sound Art*. 1 edition. New York: Bloomsbury Academic.

Zareei, Mo H., Dale A. Carnegie, and Ajay Kapur. 2015. “Physical Glitch Music: A Brutalist Noise Ensemble.” *Leonardo Music Journal*, September 21, 63–67.

Zimoun. 2013. “329 prepared dc-motors, cotton balls, toluene tank” Online at: www.zimoun.net/2013-329.html. Accessed September 2015.

Zimoun. 2014. “157 prepared dc-motors, cotton balls, cardboard boxes 60x20x20cm.” Online at: www.zimoun.net/2014-157.html Accessed September 2015.

CRACEA: GESTURAL PERFORMANCE

Mary Mainsbridge
University of Technology Sydney
Mary.M.Mainsbridge@student.uts.edu.au

ABSTRACT

This paper outlines the concepts and aims behind the musical performance piece, *Cracea*, composed for bass, drums and gestural instrument, the Telechord. The work forms part of practice-based research into the design and application of gestural systems for musical performance. It also provides a platform to assess the influence of gestural control on my own performance practice as a vocalist and live electronic musician and for developing an understanding of the broader themes concerned with gestural interaction in music.

In gestural performance, the absence of a tangible object requires the performer to develop their unique physical language and refine their movement skills to achieve fine-grained musical control. Drawing on research by movement theorists in dance and embodied interaction design, I explore how gestural systems can contribute to enhanced bodily and kinaesthetic awareness for musicians, facilitating self-discovery and the exploration of new movement patterns.

1. INTRODUCTION

Cracea is the second in a series of six works written for electronic percussion, electric bass and movement-controlled instrument, the Telechord. The virtual instrument links the body's proportions and the ratios between the limbs, torso and head to control hybrid physical models and construct four-part harmonies. Projected visuals amplify the performer's movements, depicting the relationships between key joints as strings or wires. The semi-improvised piece has been performed at the 2014 Sydney Fringe Festival and at the 2014 and 2015 Beams Arts Festival in Sydney.

The Telechord's design incorporates whole body interaction and visual cues to guide gestural interaction for the performer and aid audience comprehension. Although the work is a musical piece, it also crosses over into the categories of instrument, composition and software system, reflecting the overlapping functions that often characterise current digital musical systems.

The primary aim of the work is to stimulate awareness of ingrained bodily habits by encouraging movement awareness through exploration and improvisation. Relying on computer vision to detect continuous motion data, the Telechord utilises remote sensing to minimise behavioural constraints on the performer. Designing an open, adaptable system that translates uninhibited human movement into audio processes is a major goal of this artistic and practice-based research.

2. RELATED WORK

The embodied interaction design approach that frames this research emphasises the value of felt experience in informing decision making throughout gestural interface prototyping for performance applications. This approach draws on the theory of embodied cognition, which highlights the primacy of bodily experiences in shaping thought. Marc Leman (2008) extends this theory to the musical sphere, introducing the term embodied music cognition to affirm the centrality of the body in our experience of music.

Much work has been undertaken in the related fields of audio-visual performance and gestural augmentation of vocal performance that complements the emerging acknowledgement of the multimodal and embodied nature of human perception and cognition. Of particular relevance are pieces that incorporate non-tactile systems in digital musical instrument design, sound installations and dance contexts. Myron Krueger's pioneering installation work, VIDEOPLACE, incorporates a vision-based system that tracks hand, finger and whole body motions to allow individuals to manipulate graphic objects with unencumbered gestures (Krueger 1983). A two-dimensional virtual reality world is created using cameras and projectors to enable multiple users, represented as silhouettes, to interact with digital objects, using a rich set of gestures.

The Very Nervous System (VNS) by David Rokeby also provides a responsive environment that transforms the body of the performer or installation visitor into an instrument through a combination of video cameras, image processors, synthesisers and a sound system. The work strives to create an intimate experience within space scaled to the human form to balance the logical detachment of the computer (Rokeby 2010). Similar ideas are explored in the collaborative solo performance work, *Pikapika*, which seeks to reintegrate physicality into technological music and dance performance (Bahn, Hahn & Trueman 2001). The voice assumes a primary role in guiding improvisations and informing the design of composed instruments and mappings between physical actions and sound.

The voice is considered the body's original and most intimate instrument, bearing the emotional expressive imprint of the performer (Emmerson 2007). This individual body signature is further magnified when the voice intersects with movement in a motion-controlled interface. The mutual influence of movement and the voice has inspired a diverse range of gesturally augmented vocal systems, including Laetita Sonami's Lady's Glove (Bongers 2000); Michel Waisvisz's The Hands (Waisvisz 1985); Donna Hewitt's eMic (Hewitt 2006) and Sidney Fel's Grassp system (Fels, Pritchard & Vatikiotis-Bateson 2009). These works highlight the range of creative applications for extending the natural capacity of their voice through gestures employed during vocal performance and speech.

3. PERFORMANCE OVERVIEW

Cracea marked a transition in my practice from using a gestural controller that tracks upper body motions to manipulate digital audio effects and trigger MIDI note events to performing with an instrument that encourages improvised and spontaneous movements through whole body interaction. In the piece I controlled the pitch and timbre of hybrid physical models according to the velocity and distance between key joints, indicated in Figure 1. The piece explored notions of expansion and contraction of the limbs in relation to the torso to influence harmonic intervals. Fully extended limbs produced large intervals of octaves, while more contained postures produced major and minor thirds, and perfect fourths and fifths.

I began the performance with a solo improvisation featuring a distorted string sound, altering distortion levels with horizontal head movements. The introduction was intended to establish a clear mapping between sound and action for the audience before the entrance of the rhythm section, which grounded the remainder of the piece in forceful, punctuated percussion.

As the piece progressed, the voice acted as an exciter that shaped the overall tone of each sound. Using this input mode, I could combine the two variables of voice and movement to alter the original physical model, adapting my usual playing style as I sought to co-ordinate both elements.

In the absence of tangible feedback, visual feedback and an interaction metaphor were implemented as design strategies to enhance the nuances available through the gestural system. A string instrument metaphor was used to structure the performance experience. Leon Theremin also relied on a string metaphor to describe the design of the theremin, referring to the hand's distance from an electrode in which "the invisible string could be played by changing the hand from 50 to 10 cm or else through moving the wrist and fingers" (Theremin 1999). Adopting this metaphor became an invaluable aid to guide my movements when performing with the Telechord publicly for the first time.

Visual feedback projected onto a scrim in front of the performers depicted joint position data and the amplitude of the audio signal. It also amplified the geometry of the body's evolving postures, represented in the kinesphere, defined by Rudolf Laban (1988) as the space immediately surrounding the body. This abstracted visualisation was designed to display the virtual strings connecting selected body joints and to provide additional feedback aimed at assisting the performer to calibrate their actions spatially. Both the string metaphor and visual feedback were intended to nurture new forms of musicianship that encourage movement awareness in gestural performance.

4. THE TELECHORD

The Telechord integrates real-time physical modelling synthesis and video generation derived from improvised body postures and movement patterns. The instrument enables pitch control and harmonic layering of physical models. As shown in figure 1, four virtual strings are stretched across the apex joints of the skeleton. As the body moves, the length of the virtual resonating wires surrounding the body change the pitch of four separate tones. The shifting relationships between selected joint positions produce harmonies that reflect the architecture of the human body.

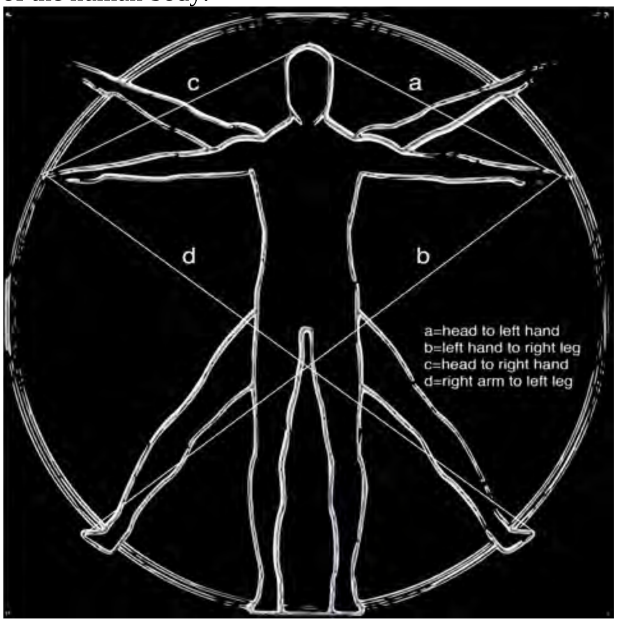


Figure 1. Virtual string positions for the Telechord.

The Telechord can perform several distinct roles in the performances, including functioning as an accompaniment for the voice, augmenting the voice with a range of physical models, and as a solo and ensemble instrument. The system also offers a framework for discovering and experimenting with sound-movement relationships and exploring individual movement tendencies and patterns.

4.1. Audio System

Using IRCAM's Modalys[1] software to simulate a range of vibrating bodies, termed 'objects', a palette of sounds were designed to suit each of the sections in the work. Modalys facilitates virtual instrument design by offering a variety of materials such as plates, strings and

[1] http://forumnet.ircam.fr/product/modalys-en/

membranes, to be matched with exciters including bows and hammers. Using several examples provided with the program as a starting point, alterations were made with LISP before each patch was exported into Max/MSP[1].

The performer has the ability to navigate between separate materials and exciters, morphing between metallic, string and membrane sounds. SimpleKinect[2] middleware is used to send OSC information from the Kinect to the Max/MSP patch transferred from Modalys. As well as detecting position information, the patch also couples acceleration to the excitation of the sound, enabling the performer to gain more nuance from a particular note by varying the speed and intensity of their movements.

The potential of physical models to provide a sonic environment from which to explore gesture has attracted much interest from researchers (Leman et al. 2008). In *Real Sound Synthesis for Interactive Applications* Perry R. Cook writes, “Our evolution and experience in the world has trained us to expect certain sonic responses from certain input gestures and parameters” (Cook 2002). Incorporating these expectations into interaction design potentially offers a familiar grounding drawn from understandings of the physical world.

4.2. Visual Feedback

The visual feedback for the instrument is deliberately simple, comprising lines that visualise the vibrating strings between the performer’s joints. These strings change colour in response to the speed and magnitude of the performer movements. Additionally, a particle system displays the amplitude of the instrument’s audio output.

The feedback helps to illuminate the mapping for the performer and audience by accentuating the lines and shapes the body forms in relation to the sounds, illustrating its underlying sound-movement mappings.

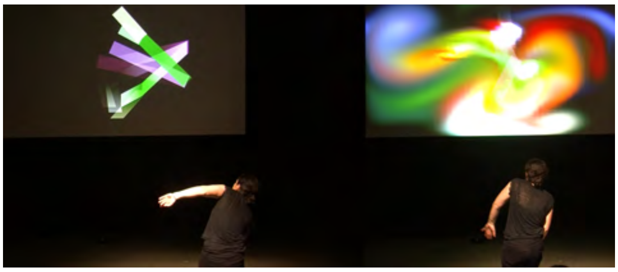


Figure 2. Telechord visualisations.

As the system is always ‘on’, it is also important for the performer to be able to identify which string is activated so that they can adequately control each separate note of the chord. For this reason the visual feedback is projected at the same scale as the performer’s body. The performer sees the projection directly before them through a scrim, yet can still maintain eye contact with the audience.

[1] https://cycling74.com/products/max/
[2] http://deecerecords.com/kinect/

5. KINAESTHETIC AWARENESS

Nurturing kinaesthetic awareness and proprioceptive skills emerged as a core aspect of this practice-based research. Aaron Levisohn (2007) proposes to expand our understanding of the body not only as a controller of multimedia systems but also “as a unique medium unto itself.”

Cracea draws on similar approaches to dance-based and virtual reality systems that are created purposely to develop proprioceptive skills, including Thecla Schiphorst and Susan Kozel’s work, *whisper: wearable body architectures* (2002), and Levisohn’s (2007) mixed reality system that enables users to experience hand movement in a different way by altering visual and aural perception. Like performer/designers who embrace embodied interaction design such as Danielle Wilde (2011) and Jin Moen (2006), I am interested in developing interaction environments where movement can be explored. This design focus encourages a holistic approach to playing a gestural instrument by creating an increased awareness of the movement possibilities available to performers.

The Telechord was initially designed to support a broad range of movement, so that tension is not held in the body through repetitive actions associated with playing traditional acoustic instruments. In musical training, the instrument is often considered superior to the needs of the performer’s body. The musician is taught to minimise their movements and become almost transparent to make the instrument ‘sing’ (McClary 1995). To avoid this type of disembodiment, I deliberately focused on strengthening an awareness of my personal movement patterns. I favoured a design that promoted variations in posture and balanced exertion evenly throughout the body’s frame.

The movement explorations encouraged through the design reflect a desire for balance and a sense of self control and body mastery, a valuable skill for musicians wishing to gain proficiency in gestural performance. This process is reminiscent of somatic practices that encourage bodily awareness, balance and conditioning, such as yoga, Feldenkrais and Alexander Technique.

6. REFLECTION

During rehearsals and performances of *Cracea*, I found myself progressively moving beyond my usual movement range and ingrained habits, occasionally crossing into the dance realm. However without formal dance or movement training, I found it challenging to achieve consistent and repeatable sonic results, particularly in the areas of pitch control and harmony. To help calibrate my movements and execute smoother and more accurate pitch transitions, I often visualised the instrument as a physical entity, at times mimicking the expressive circular motions a guitarist makes with their instrument or imagining bowing or plucking the Telechord’s virtual strings.

For iconic performers of ‘air’ instruments like Clara Rockmore, the importance of performing with visual imagery is an essential part of effective technique. Rockmore drew on a string metaphor to guide performance with the theremin. Similarly, the string metaphor underpinning the Telechord's design helped to structure my performance as I mentally referenced virtual objects to produce more subtle and precise control. Even though the visual feedback illustrated this metaphor through lines connecting active joints, I found myself relying more on an internal mental image of the virtual strings in association with audio feedback to structure my experience.

This technique highlights a two-pronged approach endorsed by somaesthetician, Richard Shusterman (2009), integrating unreflective spontaneous performance with reflective bodily awareness. The problem of using spontaneous movements alone in gestural performance is that they are a product of acquired habits, and without conscious reflection some of these habits can prove damaging. Attending to our movement patterns using body awareness techniques drawn from somatic disciplines and engaging in conscious reflection may contribute to addressing unwanted movement habits. Performing with the Telechord enabled me to engage in a similar process that has improved my sensorimotor skills, awareness of my body during performance, and a willingness to experiment with new movement behaviours to achieve more diverse sonic outcomes.

Perhaps a greater focus on the somatic dimension, as it relates to performance technique and learning, can increase the relevance of new gestural instruments beyond academic and engineering communities. Uniting technical design priorities with attention to somatic aspects that address how the instrument fits the body could thus make gestural systems more relevant to musicians from a broader range of backgrounds (Paine 2015).

7. CONCLUSION

Contributing to existing research on whole body interaction, the piece *Cracea*, written for the Telechord, explored ways to create direct and transparent mappings linking harmonic intervals to the architectural proportions of the body. The string metaphor underpinning the instrument's design informed movement-based control of hybrid physical models activated by voice and movement. Performing the work offered opportunities to explore techniques aimed at encouraging the development of kinaesthetic skills.

First-hand insights throughout the design and rehearsal process also revealed the instrument's influence on expanding movement awareness and range, highlighting the importance of direct engagement with gestural systems in improving movement mastery among musicians with no formal movement training. Through these explorations I developed new ways of moving by firstly observing and then stepping beyond my usual physical patterns to achieve more nuanced expression during movement-based performance.

8. REFERENCES

Bahn, C., Hahn, T. and D. Trueman. 2001. "Physicality and Feedback: A Focus on the Body in the Performance of Electronic Music." *Proceedings of the International Computer Music Conference*, pp. 44-51.

Bongers, B. 2000. "Physical Interfaces in the Electronic Arts." In M.M. Wanderley and M. Battier (eds.) *Trends in Gestural Control of Music*, Paris: Ircam, pp. 41-70.

Cook, P. 2002. *Real Sound Synthesis for Interactive Applications*. MA, US: A K Peters Limited.

Emmerson, S. 2007. *Living Electronic Music*. Aldershot, UK: Ashgate Publishing.

Fels, S., Pritchard, R. and E. Vatikiotis-Bateson. 2009. "Building a Portable, Gesture-to-Audio/Visual Speech System." *Proceedings of the International Conference on Audio-Visual Speech Processing*, pp. 13-18.

Hewitt, D. G. 2006. "Compositions for Voice and Technology." PhD Thesis, University of Western Sydney.

Krueger, M. W. 1983. *Artificial Reality.* MA, US: Addison-Wesley.

Laban, R. 1988. *The Mastery of Movement*. Plymouth, UK: Northcote House.

Leman, M., F. Stynes and N. Bernadini. 2008. "Sound, Sense and Music Mediation: A Historical-philosophical Perspective." In P. Polotti and D. Roccesso (eds.) *Sound to Sense, Sense to Sound: a State of the Art in Sound and Music Computing*, Berlin: Logos Verlag, pp. 15-46.

Levisohn, A. 2007. "The Body as Medium: Reassessing the Role of Kinaesthetic Awareness in Interactive Applications." *Proceedings of Multimedia '07*, pp. 485-488.

McClary, S. 1995. "Music, the Pythagoreans, and the Body." In S. I. Foster (ed), *Choreographing History*. Bloomington: Indiana University Press, pp. 82-105.

Moen, J. 2006. "KinAesthetic Movement Interaction: Designing for the Pleasure of Motion." PhD Thesis, University of Stockholm.

Paine, G. 2015. "Interaction as Material: The Techno-somatic Dimension." *Organised Sound,* 20(1): 82-89.

Rokeby, D. 2010. "Works: Very Nervous System (1986-1990)." Available online at davidrokeby.com/vns.html. Accessed November 2015.

Schiphorst, T. and K. Andersen. "Between Bodies: Using Experience Modeling to Create Gestural Protocols for Physiological Data Transfer." *Proceedings of the CHI 2004 Fringe*, pp. 1-8.

Shusterman, R. 2009. "Body Consciousness and Performance: Somaesthetics East and West." *The Journal of Aesthetics and Art Criticism*, 61(2): 133-145.

Theremin, L. 1999. "Recollections." *Contemporary Music Review,* 18(3): 5-8.

Waiswisz, M. 1985. "The Hands: A Set of Remote MIDI Controllers, *Proceedings of the International Computer Music Conference*, pp. 313-318.

Wilde, D. 2011. "Swing That Thing: The Poetics of Embodied Engagement." PhD Thesis, Monash University.

MUSIC OF 18 PERFORMANCES: EVALUATING APPS AND AGENTS WITH FREE IMPROVISATION

Charles Martin
Research School of Computer Science
Australian National University
charles.martin@anu.edu.au

Henry Gardner
Research School of Computer Science
Australian National University
henry.gardner@anu.edu.au

Ben Swift
Research School of Computer Science
Australian National University
ben.swift@anu.edu.au

Michael Martin
Research School of Finance, Actuarial Studies & Applied Statistics
Australian National University
michael.martin@anu.edu.au

ABSTRACT

We present a study where a small group of experienced iPad musicians evaluated a system of three musical touch-screen apps and two server-based agents over 18 controlled improvisations. The performers' perspectives were recorded through surveys, interviews, and interaction data. Our agent classifies the touch gestures of the performers and identifies new sections in the improvisations while a control agent returns similar messages sourced from a statistical model. The three touch-screen apps respond according to design paradigms of reward, support, and disruption. In this study of an ongoing musical practice, significant effects were observed due to the apps' interfaces and how they respond to agent interactions. The "reward" app received the highest ratings. The results were used to iterate the app designs for later performances.

1. INTRODUCTION

This paper describes the evaluation of a system of touch-screen musical instrument apps and server-based computational agents in a controlled study of 18 free-improvised performances. Free-improvised ensemble music is performed without any plan for the performance and our system is designed to react to the performance structure that emerges while a group is playing. Improvisations can be considered to be segmented by new musical ideas (Stenström 2009, pp. 58–59) and our ensemble-tracking agent searches for these new ideas by classifying and analysing performers' touch gestures. Three musical apps have been developed for the Apple iPad platform that receive messages from this agent and react by updating their interfaces in real-time. Each of the three apps encodes a different behavioural model of interaction in ensemble improvisation. A "reward" model gives performers access to new notes at each new section of the performance, a "disruption" model interrupts performers who stay on one gesture for too long, and a "support" model plays complementary sounds when performers focus on individual gestures.

A group of three touch-screen musicians with more than a year of performance experience with the apps were participants in the study. While concert experience had suggested that the ensemble-tracking agent interacted with the group accurately and could enhance improvisation, a formal experiment was conducted to evaluate the agent system under controlled conditions and compare the three apps. To assess the accuracy of the ensemble-tracking agent, a control agent was developed that generates similar messages randomly from a statistical model. In a methodology that combined a balanced experimental design with rehearsal processes, the group performed a series of 18 improvisations on all combinations of the three iPad interfaces and the two agents. We performed quantitative analyses of survey ratings from the musicians, and on the number of agent messages sent during performances, as well as qualitative analysis compiled from interviews.

The results support the effectiveness of our ensemble-tracking agent, although the source of agent interventions was seen as less important than how the apps responded. The app condition was found to have a significant main effect on the performer's responses to several questions, including the quality and level of creativity in performances. The app featuring the "reward" model showed the most positive response with the performers actively seeking out interaction with the agent when using this app. The performers articulated problems with the other two apps while still finding ways to use them in interesting improvisations and their responses were used to redesign the apps for later performances. Following a review of prior work in this field, in Section 2 we will describe the construction of our

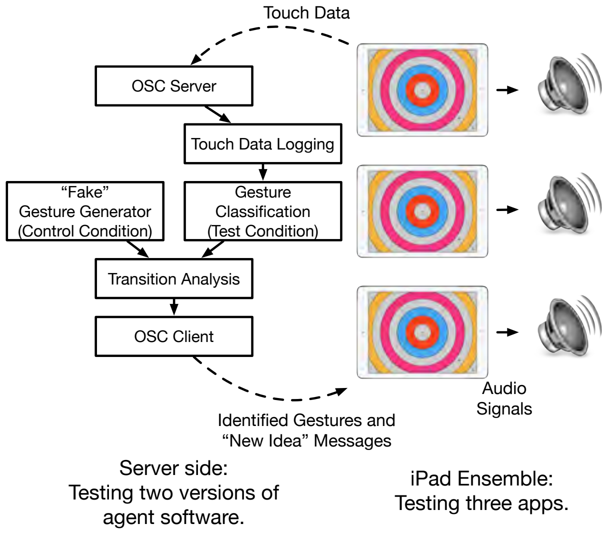


Figure 1. A system diagram of our agent software interacting with our iPad instruments. In the test condition, touch messages are classified as gestures by a Random Forest classifier, while in the control, gestures are generated from a statistical model disconnected from the performers' current actions.

system of apps and agents, Section 3 will describe our experimental design, and our results will be analysed and discussed in Section 4.

1.1 Related Work

The "Laptop Orchestra" (Bukvic et al. 2010; Trueman 2007) (LO), where multiple performers use identical hardware and software in musical performance, is now established in computer music practice and has an expanding compositional repertoire (Smallwood et al. 2008) and pedagogy (Wang et al. 2008b). These ensembles often use artificial intelligence agents as ensemble members (Martin et al. 2011) or as a "conductor" (Trueman 2007) to provide cohesive direction of broad musical intentions. In our study, two fundamental designs of these mediating agents are evaluated: one using a statistical Markov model (Ames 1989), and one using machine learning algorithms to follow the performers (Fiebrink et al. 2009).

With the emergence of powerful mobile devices such as smartphones and tablets, ensembles of "mobile music" (Gaye et al. 2006; Jenkins 2012) performers have appeared, taking advantage of the many sensors, touch screen interfaces, and convenient form-factors of these devices (Tanaka 2010). These ensembles have used phones to perform gamelan-like sounds (Schiemer and Havryliv 2007), sensor-based music (Wang et al. 2008a) or explore touch interfaces (Oh et al. 2010). Both smartphones (Swift 2013) and tablets (Martin et al. 2014) have been used in improvising ensembles and Williams (2014) has noted their utility in exploratory, collaborative music making. While mobile instruments have often been aimed towards beginners (Wang 2014; Wang et al. 2011), we have been developing a long term musical practice by experienced performers.

There are a wide range of approaches for evaluating

#	Code	Description	Group
0	N	Nothing	0
1	FT	Fast Tapping	1
2	ST	Slow Tapping	1
3	FS	Fast Swiping	2
4	FSA	Accelerating Fast Swiping	2
5	VSS	Very Slow Swirling	3
6	BS	Big Swirling	3
7	SS	Small Swirling	3
8	C	Combination of Swirls and Taps	4

Table 1. Touch-screen gestures that our classifier is trained to identify during performances. When gestures are summarised in transition matrices, the gesture groups are used instead, producing 5×5 matrices.

new digital musical instruments, but it is generally accepted that the performer is the most important stakeholder (O'Modhrain 2011), particularly when performing improvised music. Gurevich et al. (2012) have used a grounded-theory approach to identify styles and skills that emerge when multiple participants engage with very simple electronic instruments. Fiebrink et al. (2011) asked users to repeatedly evaluate interactive musical systems that use machine learning across a number of novel criteria, and this "direct evaluation" was found to have more utility than a typical cross-validation approach for machine learning systems. A long-term ethnographic study of the Reactable table-top surface observed collaborative and constructive processes (Xambó et al. 2013) in video footage of improvised performances. Ethnographic techniques have also been used over natural rehearsal and development processes such as for Unander-Scharin et al.'s (2014) "Vocal Chorder", where an autobiographical design process transitioned into an interface developed for other performers. Our study uses a reheasal-as-research methodology where multiple performances are evaluated in a single session through short surveys and interviews. As our iPad ensemble had already established a performance practice (Martin 2014) they were able to test the six experimental conditions with 18 improvisations in one session, an unprecedented number in musical interface evaluation.

2. SYSTEM DESIGN

The following sections detail the construction of our iPad apps, ensemble-tracking agent, and control agents. An overview of the system architecture is given in Figure 1 which shows the two important parts of our agent software: a gesture classification system which uses machine learning algorithms to identify each performer's actions, and a performance tracking system which analyses the ensemble's transitions between touch screen gestures to identify important "new-idea" moments in these performances. We also describe a "fake" gesture generator, used as an experimental control, where gestures were generated randomly from a statistical model derived from a live performance. The gesture generator takes the place of the classifier while other parts of the system remain the same.

Our iOS iPad apps are developed in Objective-C, with the `libpd` library used for audio synthesis. Our server

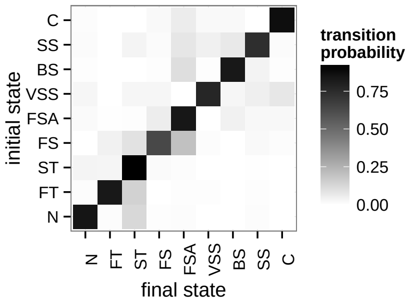


Figure 2. Plot of the transition matrix used for the Markov model in the generative agent.

software is developed in Python and communicates with the iPad apps over a WiFi network connection using the OSC message format (Freed and Schmeder 2009). The iPad apps send messages to the server for each touch-screen event and the server logs all of this touch information for gesture classification and for later analysis.

2.1 Gesture Classifier

Previous work has identified a vocabulary of gestures used by expert percussionists on touch-screen interfaces (Martin et al. 2014). Our agent is able to identify nine of these touch-screen gestures (see Table 1) using each performer's touch-data at a rate of once per second. The server records these gestures and also sends them back to the performers' iPads.

Classification is accomplished using a Random Forest classifier algorithm (Breiman 2001) provided by Python's `scikit-learn` package. This was trained using examples of the touch-screen gestures recorded in a studio session by our app designer. The input to the classifier is a feature vector of descriptive statistics from the last five seconds of each performer's touch data. The timing parameters for our classifier were tuned by trial and error and previous research (Martin et al. 2015) has shown that our classifier has a mean accuracy of 0.942 with standard deviation 0.032 under cross-validation, comparable to other systems that recognise command gestures (Wobbrock et al. 2007).

2.2 Generating Fake Gestures

In order to evaluate the effect of our gesture classifying agent (CLA) on performances in our experiment, we developed a contrasting system that generates fake gestures (GEN) to be used as a control. As the rest of our agent software remains the same (see Figure 1), the fake gestures and fake "new-idea" messages would be recorded and reported back to the iPads in the same way as with the ensemble-tracking agent.

To build this control agent, a live touch-screen performance of the iPad ensemble was analysed with our classification system and the resulting sequence of states was used to construct a first-order Markov model. The concept of using a Markov model to generate data is a common design pattern in computer music and Ames (1989) has described how it can be used to algorithmically compose melodies or vary other musical parameters. In our case, the model was used to generate fake gesture classifications similar to the gestural output of our touch-screen ensemble. As it is statistically similar to the changes induced by the classifying agent, but decoupled from the performers' actual gestures, our generative agent was used as a control in our experiment to expose the effect of an intelligent mediation of live performance.

2.3 Transitions, Flux, and New Ideas

Our classifying agent is designed to identify the musical sections present in improvised musical performances and pass this information to the iPad interfaces operated by the performers. A "new-idea" in our system is defined as a moment in the performance where, for the whole ensemble, transitions between different gestures increases sharply over 30 seconds. The implementation of the system is more fully explained in previous research (Martin et al. 2015), but is presented here in brief.

An improvised musical performance can be modelled as a sequence of abstract musical gestures for each performer in the ensemble. In the present study, these gestures are either identified by the gesture classifier (CLA) or generated by our statistical model (GEN). Transitions between different gestures over a certain window of time can be summarised by a transition matrix P constructed in a similar way to the transition matrix of a Markov chain (Swift et al. 2014).

The matrices for each performer can be averaged to summarise the whole ensemble's transition activity. Our agent software compares transition matrices by applying a matrix measure, flux, which is defined as follows:

$$\text{flux}(P) = \frac{\|P\|_1 - \|\text{diag}(P)\|}{\|P\|_1} \tag{1}$$

where $\|P\|_1 = \sum_{i,j} |p_{ij}|$ is the element-wise 1-norm of the matrix P and $\text{diag}(P)$ is the vector of the main diagonal entries of P.

The flux measure is equal to 0 when all the non-zero elements of the transition matrix P are on the main diagonal, that is, when performers never change gesture. The measure will be equal to 1 when no performer stays on the same gesture for two subsequent classifications.

In our agent software, the flux of the ensemble is calculated each second for the two preceding 15 second windows of gestures reduced to their "groups" (see Table 1). If the flux of the ensemble has increased over these windows by a certain threshold, the system sends a new-idea message to the performers's iPads. The iPad apps include a rate-limiting function that prevents them reacting to several measurements of the same new-idea event by ignoring messages for at least 10 seconds after responding to a new-idea. The timing parameters and threshold for detecting new-ideas were tuned by trial and error within the research group. As well as reporting new-idea events, our agent sends the entire gesture classification sequence to the apps, which are able to respond to long sequences of identical gestures (old ideas) as well as the new-idea messages.

2.4 iPad Apps

Three different iPad apps were chosen from a repertoire

Figure 3. The three apps used in this study, from left to right: Bird's Nest (BN), Singing Bowls (SB), and Snow Music (SM).

of six apps created by our designer and routinely used by our ensemble. The three apps, Bird's Nest (BN), Singing Bowls (SB), and Snow Music (SM), are shown in Figure 3. Each app features a free-form touch area where tapping produces short sounds and swiping or swirling produces continuous sounds with volume controlled by the velocity of the moving touch point. While these apps share a paradigm for mapping touch to sound, their different sound material and contrasting designs for interaction with our agents make them three distinct instruments.

Bird's Nest and Snow Music present nature-inspired interfaces to the performers. In Bird's Nest, performers create a soundscape from a northern Swedish forest with bird samples, field recordings, percussive sounds, and a backdrop of images collected at that location. Snow Music emulates a bowl of amplified snow, where performers manipulate field recordings of snow being squished, smeared, stomped and smashed. Singing Bowls presents users with an annular interface for performing with percussive samples. Rings on the screen indicate where touches will activate different pitches of a single-pitched percussion sample. Tapping a ring will activate a single note while swirling on a ring will create a sustained sound reminiscent of that of Tibetan singing bowls.

The apps' response to messages from the agent followed three distinct paradigms. Bird's Nest was designed to **disrupt** the musicians' performance. Based on gesture feedback from the agent, the app would watch for runs of identical gestures and then switch on looping and autoplay features in the user interface in order to prompt new actions by the performers. New-idea messages were used to randomise the sounds available to the user from a palette of sample and pitched material.

Snow Music used a **supportive** paradigm. The app would watch for sequences of similar gestures and activate extra layers of complementary sounds. For instance, the app would support a run of tapped snow sounds by layering the taps with glockenspiel notes while a backdrop of generative bell melodies would be layered on top of the sound of continuous swirling gestures. When the performer moves on to other gestures, the supportive sounds were switched off. New-idea messages in Snow Music changed the pitches of the supportive sounds and the snow samples available to the performer. While the actions of the supportive sounds were shown on the screen, the performers were not able to control them directly.

Set	Perf. 1	Perf. 2	Perf. 3
0	orientation		
1	SM, CLA	BN, GEN	SB, CLA
2	BN, CLA	SB, GEN	SM, GEN
3	SB, CLA	SM, CLA	BN, GEN
4	SB, GEN	BN, CLA	SM, GEN
5	BN, GEN	SM, CLA	SB, CLA
6	SM, GEN	SB, GEN	BN, CLA
7	interview		

Table 2. The experiment schedule showing the balanced ordering of apps and agents. The experiment was performed in one session divided by breaks into six groups of three five minute performances.

Finally, the Singing Bowls app **rewarded** the player's exploration of gestures with new pitches and harmonic material. This app only allows the performer to play a limited number of pitches at a time. When the ensemble's performance generates a new-idea message, the app rewards the players by changing the number and pitches of rings on the screen. The pitches are taken from a sequence of scales so that as the performers explore different gestures together, they experience a sense of harmonic progression.

3. EXPERIMENT

Our experiment took the form of a lab-based study under controlled conditions. Although analogous to a rehearsal call for professional musicians in its length and artistic intent—a performance of this ensemble actually took place some four weeks later at an art exhibition—the research intent of this experiment meant that it was quite an unusual rehearsal from the musicians' perspective.

In the experiment, two agents (a classifying agent: CLA, and a generative agent: GEN) were crossed with three iPad apps (Bird's Nest: BN, Singing Bowls: SB, and Snow Music: SM) to obtain the six independent conditions. The

Figure 4. The ensemble setup for the lab study shown from one of two camera angles. Each performer's sound was dispersed through a large loudspeaker directly behind them and simultaneously recorded.

ensemble were asked to perform improvisations limited to five minutes each and to immediately fill out questionnaires after each improvisation. It was determined that 18 of these sessions would fit into a standard three-hour rehearsal session which allowed for three trials of each of the six independent conditions.

The entire rehearsal was divided into six sets of three performances (see Table 2) preceded by an orientation and followed by an open-ended interview. In each set, the musicians used each app once and the order of apps was permuted between sets in a balanced design following Williams (1949) to offset local learning effects. Successive performances with each app alternated between the two agents. The experiment was blinded insofar as the performers were aware that two agents were under investigation but were not made aware of the difference between them or of which agent was used in each performance.

The experiment took place in an acoustically treated recording studio (see Figure 4). The performers were seated in the recording room while the two experimenters were present in a separate control room. The experiment was video recorded with two angles[1] which allowed the performers' faces and screens to be seen. The sound of each iPad was recorded from the headphone output in multitrack recording software[2] and simultaneously diffused through large monitor speakers behind the performers. Audio from a microphone directly in front of the ensemble as well as from a microphone in front of the experimenter was also recorded to capture discussion during the experiment and during the post-session interview. In each performance session all touch-interaction messages from the three performers' iPads were recorded (even though only the CLA agent made use of this information), as were the messages returned to the performers by the agents.

3.1 Participants

The participants in this study (Performer A, Performer B, and Performer C) are members of an ensemble established to perform improvised music with the apps and agents under investigation as well as acoustic percussion instruments. All three participants are professional percussionists and had worked together previously in educational and professional contexts. The fourth member of this ensemble (Experimenter A) was also the designer of the apps and agents but did not participate in the performances in this study. A second researcher (Experimenter B) assisted with running the study. The two experimenters are also experienced musicians.

Over the 14 months prior to this study, the performers had engaged in a longitudinal process of rehearsals and performances parallel to the development of the apps and agent. The process and other results of this longitudinal study have previously been reported (Martin 2014; Martin and Gardner 2015; Martin et al. 2014).

The three performers were chosen to participate in the present study due to their high level of skill and experience in iPad performance and their capacity for self-evaluation. Cahn (2005, pp. 37–38) has written about the strong learning effect present in new improvisation ensembles, where members overcome initial inhibitions to test the limits of newfound musical freedom with "severe departures from normal music making". This phase is followed by a plateau of thoughtful free-improvisation where "listening and playing come into more of a balance". The significant experience by the performers in this study meant that all of the performances recorded had the potential to be of high quality.

3.2 Questionnaires

At the end of each performance, the performers filled out written surveys consisting of the following questions on a five point Likert-style scale (Very Bad, Bad, Neutral, Good, Excellent). The two experimenters present during the lab study were also surveyed on Question 1.

1. How would you rate that performance?
2. How would you rate the level of creativity in that performance?
3. How did the agent's impact compare to having it switched off?
4. How well were you able to respond to the app's actions?
5. How well were you able to respond to the other players' actions?
6. How was the app's influence on your own playing?
7. How was the app's influence on the group performance?

The format and content of our questionnaire follows other evaluations of improvised performance, including Eisenberg and Thompson (2003), in evaluating overall quality, creativity, and ensemble interaction, however we added specific questions to evaluate the overall impact of the agents and the changes that they caused in the apps.

4. RESULTS

In the following sections we analyse and discuss the data collected in the study session. This corpus consists of 57 minutes of interviews, 92 minutes of performances,

[1]The video recorders were a GoPro HERO 3+ Black and a Zoom Q2HD both set to record in 1920*1080 resolution.

[2]The audio recording was made through a Presonus Firepod interface in Apple Logic Studio.

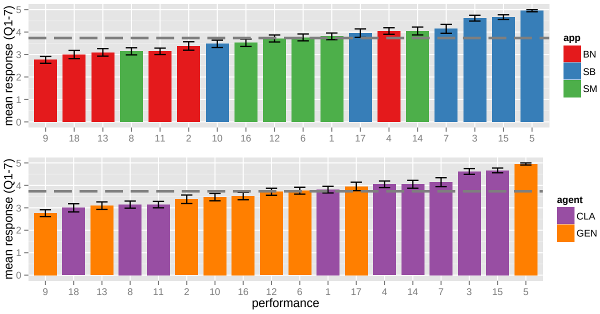


Figure 5. Performances ordered by the mean response to all questions. The grand mean is shown as a dashed horizontal line. The distribution of apps is striking with SB eliciting the highest ratings. Two thirds of the CLA agent performances appear above the mean.

32.2MB of touch and interaction data as well as the experimenters' notes. We will first discuss the data from surveys and agent-app interaction before considering the interview responses.

4.1 Survey Data

The survey responses from each question were analysed separately using univariate two-way repeated measures analysis of variance (ANOVA) procedures to determine the significance of main and interaction effects of the two indepedent variables (app and agent). Post-hoc Bonferroni-corrected paired t-tests were used to assess significant variations between each of the six experimental conditions. This is a standard procedure for significance testing used in human-computer interaction studies (Lazar et al. 2010).

Results from five of the seven questions (1,2,4,6,7) were found to be significant and will be considered below in detail. The other questions (3,5) were not significantly affected by the change of apps and agents. The normal variations of musical interactions in between five minute performances may have affected these questions more than the independent variables.

4.1.1 Mean Response

Figure 5 shows the mean response to all questions, yielding a holistic overview of the results. For the apps, this figure shows that, in general, Singing Bowls was rated higher than Snow Music which was higher than Bird's Nest. For the agents, performances with the classifying agent were, in general, more highly rated than those with the generative agent, with six of the nine classifier performances appearing above the grand mean.

4.1.2 Performance Quality and Creativity

Questions 1 and 2 of the survey concerned the overall level of quality and creativity in each performance, the distribution of responses to these questions are shown as a box plot (McGill et al. 1978) in Figure 6. Considering all responses to Question 1 in the survey (including the two experimenters), the app used had a significant effect on the perception of quality in the performances, $F(2, 8) = 5.006, p < 0.05$. The main effect of the agent and the interaction effect were not found to be significant.

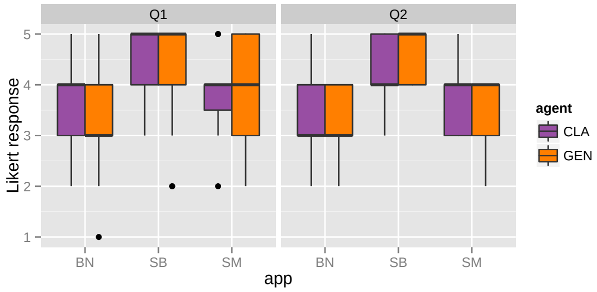


Figure 6. Distribution of performance quality (Question 1) and creativity (Question 2) ratings by app and agent. For both questions, the app had a significant effect on ratings.

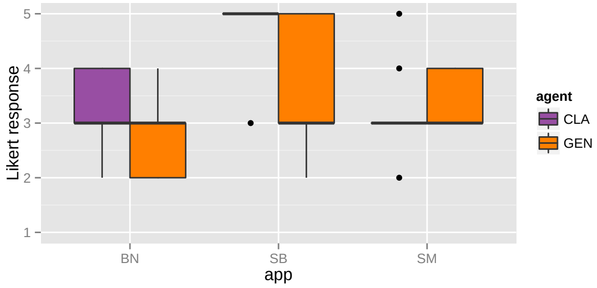


Figure 7. Distribution of ratings of individual performers' responses to the agents' actions (Question 4). The main effect of the app and an interaction effect between app and agent were found to be significant.

Bonferroni-corrected paired t-tests revealed that, without considering the agent, performances with the Singing Bowls app were of significantly higher quality than with Snow Music ($p = 0.04$) and Bird's Nest ($p = 0.002$).

A significant main effect of app was also observed on the performers' ratings of the level of creativity in their performances, $F(2, 4) = 8.699, p < 0.05$. Bonferroni-corrected paired t-tests only showed that performances of Singing Bowls with the generative agent were rated as significantly more creative ($p < 0.05$) than Snow Music with the generative agent and Bird's Nest with either agent.

4.1.3 Responding to the App's Actions

The performers were surveyed on how well they were able to respond to changes in the app interfaces caused by the agent (Question 4). Although the agent and app worked together as a system to change the interface, on the survey we called this "the app's actions" as from the performers' perspective, they were only aware of changes in the app.

A box plot of the results are shown in Figure 7. There was a significant effect of the app ($F(2, 4) = 13.32$, $p < 0.05$) and a significant interaction effect between agent and app ($F(2, 4) = 7.75$, $p < 0.05$). The effect of the agent was of borderline significance ($F(1, 2) = 16$, $p = 0.0572$). Post-hoc Bonferroni-corrected pairwise t-tests revealed that the performers were able to respond to the Singing Bowls app in combination with the classifying agent significantly better than for the other two apps with either agent ($p < 0.05$) but with only borderline significance against the Singing Bowls app with the generative

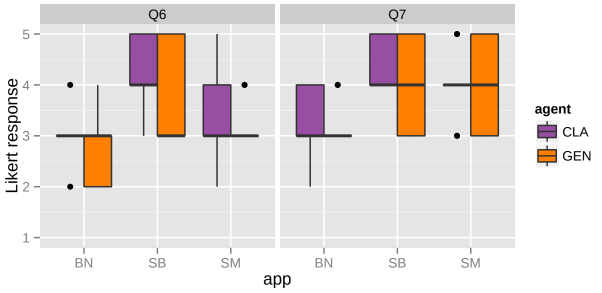


Figure 8. Distribution of responses about the influence of the app and agent on the performers' own playing (Q6), and the group's playing (Q7). In Q6, the app had a significant effect, while in Q7, the agent appears to have been more important.

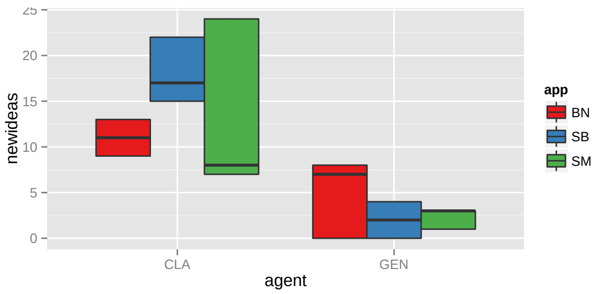


Figure 9. Distribution of the number of new-idea messages sent during performances. The GEN agent failed to match the CLA agent's behaviour, but the app also had an effect.

agent ($p = 0.11$). These tests revealed that when using the Bird's Nest and Snow Music apps and the classifying agent, performers reported that they were better able to respond to the app's actions than with the generative agent, although significance was borderline ($p < 0.1$).

4.1.4 *App/Agent Influence*

Questions 6 and 7 both relate to the influence of the app and agent on the performance with the former asking about impact on the individual and the latter on the group. By univariate ANOVA, the effect of the app on the performers' own playing was found to be significant ($F(2, 4) = 137.2$, $p < 0.01$)). The effect of the agent on the group performance was of borderline significance ($F(1, 2) = 16$, $p = 0.0572$).

A multivariate ANOVA on both outcomes showed significance only for the app's effect ($F(2, 4) = 4.238, p < 0.05$). These results suggest that although the app interface was the most important factor in the participants' perceptions of their own playing, the agent was a more important factor when considering the group.

4.1.5 *New Ideas*

As discussed in Section 2, the generative agent produced randomised gesture classifications based on a statistical distribution derived from a previous performance of the iPad ensemble. We had hoped that this agent would act as a control in our experiment by producing a similar number of new-idea messages but at times which did not correlate

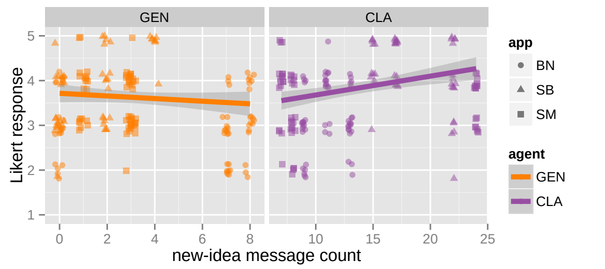


Figure 10. Number of new-idea messages in each performance plotted against performers' responses to all questions. The lines show a linear model with standard error for messages generated by the GEN and CLA agents. Responses are lower for more GEN new-ideas, but higher for more CLA new-ideas.

with activity in the live performance. However, from Figure 9, it is clear that the classifying agent produced more new-idea messages than the generative agent. We can investigate this difference by treating the number of new-idea messages as a dependent variable.

A two-way ANOVA showed that only the effect of the agent on the number of the new-idea messages was significant ($F(1, 12) = 24.19, p < 0.001$). Although the app's effect on new-ideas was not found to be significant, the number of new-ideas generated with Singing Bowls and the classifying agent appears higher than with Snow Music or Bird's Nest (Figure 9). This suggests that the musicians may have performed more creatively with Singing Bowls, cycling through numerous gestures as an ensemble.

Figure 10 shows the performers' responses to all questions against the number of new-ideas in each performance. A linear model of responses for each agent suggests that ratings decline as the generative agent produced more new-idea messages, while ratings increase as the classifying agent produced more messages. This may suggest that for the generative agent, more changes due to new-idea messages annoy the performers as they do not necessarily correspond to their actions. For the classifying agent, large numbers of new-idea messages may have been triggered by particularly creative and engaged performances which elicited higher ratings. While the generative agent did not produce the same numbers of new-idea messages as the classifying agent, if it had, the performers' responses may have been more negative.

4.2 Qualitative Analysis

Video recordings were made of the orientation briefing, the 18 performances, and the post-experiment interview. Thematic analysis (Braun and Clarke 2006) of a transcription of the interview revealed that the performers' experiences were shaped by the three apps and their interaction with the agents. This qualitative analysis was used to direct a redesign of two of the apps leading up to a concert performance of the ensemble four weeks after the experiment.

From the interview data, and confirming the analysis of the Likert data for Question 1 and 2, the performers were most satisfied with the Singing Bowls app. It was noted that Singing Bowls was "familiar as an instrument . . . responds as we're used to" (Perf. B) and that the "time went quickly in performances" (Perf. B). The ensemble noticed some performances (with the generative agent) where the Singing Bowls app "didn't change at all" (Perf. C). Rather than being discouraged, the performers tried actively to "get it to respond by copying and mimicking and getting everybody to change" (Perf. A). Because of this positive reception, Singing Bowls was deemed a success and its design was not revisited before the concert performance.

In marked contrast to Singing Bowls, performances with Snow Music felt like they "went on forever" (Perf. A), suffering from a lack of structure and motivation to keep playing with the smaller palette of snow sounds. The performers suggested that the "use of space" (i.e. silence) in Snow Music performances could improve quality and allow focus on particular sounds. The interaction with the supporting sounds was described as "lovely" (Perf. A) and "would play some really good stuff" (Perf. C). In response to these comments, design revisions were made to add to the palette of snow sounds and to refine the synthesis system to be more expressive. A sequence of harmonies was added to the pitched sounds in this app to encourage the group to continue exploring throughout the performance.

Bird's Nest performances suffered the lowest ratings from the performers who felt annoyed and "isolated" (Perf. A) by the disruptive interaction between the app and agent and found it "really hard" (Perf. C) to use creatively. While the app's sounds were "pretty" (Perf. A) it was "hard to have that flow of ideas between people" (Perf. C). It was noted that Bird's Nest was "less similar than an instrument" (Perf. B) than the other apps and that the sounds were "long. . . and the change in pitch is. . . less perceptible" (Perf. C). Following these comments, Bird's Nest was extensively revised for later concerts. The "autoplay" feature and the disruptive control of the "looping" function were removed. A sequence of images with corresponding scales and subsets of the sound palette was devised to form a compositional structure for performances. The "sounds" button was retained to refresh the palette of sounds for each scene, but, as with Singing Bowls, movement through the compositional structure depended on new-idea messages. The synthesis system for playing bird samples was refined to play sounds of varying, but usually much shorter, length.

The qualitative analysis suggested that, from the performers' point of view, the source of the agent's interventions (either responding to their gestures or generated from a model) was not as important as the way that the apps *responded* to these interventions. The "rewarding" paradigm used in Singing Bowls was the most successful in engaging the performers' interest. It was notable that with Singing Bowls the performers sought out agent interactions, particularly when the agent did not respond as was the case with the generative agent.

5. DISCUSSION

The primary limitation of the present study is the small number of participants surveyed. We surveyed only three participants with very specialised skills, so the generalisation of their responses is limited. The goal of this study was not to evaluate performances by inexperienced players but by practiced iPad musicians with an important stake in the quality of the instruments they use. We studied an expert iPad ensemble with extensive performance experience and as a result, we were able to examine more experimental conditions and more improvisations than would be feasible with beginners. As far as we are aware, no controlled studio-based study of 18 touch-screen performances by the same ensemble has been previously attempted. Given the strong preference for the Singing Bowls app, future studies with more participants may be warranted that focus only on this app to reduce the number of required trials.

The multitrack audio and video recordings of the 18 improvised performances and corresponding touch gesture data were important outcomes of this study. Other studies have used detailed logs of improvisations as a basis for analyses of keyboard (Gregorio et al. 2015; Pressing 1987) and live-coding (Swift et al. 2014) performances. We propose that performing similar analyses on our recorded performances could lead to further understanding of the structure of touch-screen improvisation and improvements in the ability of our gesture-classifying agent to track such performances.

6. CONCLUSION

Our system for ensemble touch-screen musical performance includes two server-based agents and three iPad apps. One agent classifies performers' gestures to track new ideas while the other generates similar messages from a statistical model. The three iPad apps use the responses from these agents to support, disrupt, and reward gestural exploration in collaborative improvised performances.

We have presented the results of an evaluation of this system's use in real musical performances in a formal, order-balanced study that considered surveys, interviews and interaction data by an expert iPad ensemble with 14 months of experience. The participants' high skill level allowed us to introduce a novel experimental design where a total of 18 performances over six conditions were recorded and evaluated in one rehearsal session.

Different apps were found to have a significant main effect on the performers' perception of performance quality and creativity, how well they were able to respond to interface changes, and the app's influence on individual playing. The main effect due to the agent was found only to have borderline significance on the app's influence on the group performance and the perfomers' ability to respond to interface changes. However, this question did reveal a significant interaction effect between the app and agent conditions.

While significant effects due to the agent were somewhat elusive, the study revealed that our generative agent produced significantly fewer new-idea messages than the classifying agent. Modelling the performer responses with respect to the number of new-idea messages suggests that performances with many classified new-ideas were rated highly, but frequent generated new-ideas may have had a negative impact on ratings.

The results of our study lead us to conclude that the design of an agent's interaction with a creative interface can make or break a performer's positive perception of this interaction; this design can also limit or enhance dynamic and adventurous playing. While experienced performers can create high quality, creative performances mediated by agents of many designs, connecting the agent to their actions seems to have a positive effect on how they respond to interface changes and their perception of the group performance. Rewarding users for their collaborative exploration was found to be an especially engaging paradigm for supporting creativity. The idea of disrupting performers' flow to encourage more creative interaction was roundly rejected in both quantitative and qualitative results.

This study has been a snapshot in the participants' ongoing artistic practice, and the recommendations from the performers have already been taken into account in updates to the apps for subsequent performances. Design guidelines for agent-app interaction will be further articulated in future work. While this study has concerned expert performers, given the broad interest in touch-screen computing, future investigations could consider a wider range of performers, and particularly users in musical education.

7. REFERENCES

C. Ames. The Markov process as a compositional model: A survey and tutorial. *Leonardo*, 22(2):175–187, 1989.

V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006.

L. Breiman. Random Forests. *Machine Learning*, 45(1): 5–32, 2001.

I. I. Bukvic, T. Martin, E. Standley, and M. Matthews. Introducing L2Ork: Linux Laptop Orchestra. In *Proc. NIME '10*, pages 170–173, 2010.

W. L. Cahn. *Creative Music Making*. Routledge, 2005.

J. Eisenberg and W. F. Thompson. A matter of taste: Evaluating improvised music. *Creativity Research Journal*, 15(2-3):287–296, 2003.

R. Fiebrink, D. Trueman, and P. R. Cook. A metainstrument for interactive, on-the-fly machine learning. In *Proc. NIME '09*, 2009.

R. Fiebrink, P. R. Cook, and D. Trueman. Human model evaluation in interactive supervised learning. In *Proc. CHI '11*, pages 147–156. ACM Press, 2011.

A. Freed and A. Schmeder. Features and future of Open Sound Control version 1.1 for NIME. In *Proc. NIME '09*, pages 116–120, 2009.

L. Gaye, L. E. Holmquist, F. Behrendt, and A. Tanaka. Mobile music technology: report on an emerging community. In *Proc. NIME '06*, pages 22–25, 2006.

J. Gregorio, D. S. Rosen, M. Caro, and Y. E. Kim. Descriptors for perception of quality in jazz piano improvisation. In *Proc. NIME '15*, 2015.

M. Gurevich, A. Marquez-Borbon, and P. Stapleton. Playing with constraints: Stylistic variation with a simple electronic instrument. *Computer Music Journal*, 36(1): 23–41, 2012.

M. Jenkins. *iPad Music: In the Studio and on Stage*. Taylor & Francis, 2012.

J. Lazar, J. Feng, and H. Hochheiser. *Research Methods in Human-Computer Interaction*. John Wiley & Sons, West Sussex, UK, 2010.

A. Martin, C. T. Jin, and O. Bown. A toolkit for designing interactive musical agents. In *Proc. OzCHI '11*, pages 194–197. ACM Press, 2011.

C. Martin. Making improvised music for iPad and percussion with Ensemble Metatone. In *Proceedings of the Australasian Computer Music Conference*, pages 115–118, Fitzroy, Vic, Australia, 2014. ACMA.

C. Martin and H. Gardner. That syncing feeling: Networked strategies for enabling ensemble creativity in iPad musicians. In *Proc. CreateWorld '15*, Brisbane, Australia, 2015. AUC.

C. Martin, H. Gardner, and B. Swift. Exploring percussive gesture on iPads with Ensemble Metatone. In *Proc. CHI '14*. ACM Press, 2014.

C. Martin, H. Gardner, and B. Swift. Tracking ensemble performance on touch-screens with gesture classification and transition matrices. In *Proc. NIME '15*, 2015.

R. McGill, J. W. Tukey, and W. A. Larsen. Variations of box plots. *The American Statistician*, 32(1):12–16, 1978.

J. Oh, J. Herrera, N. Bryan, L. Dahl, and G. Wang. Evolving the mobile phone orchestra. In *Proc. NIME '10*, 2010.

S. O'Modhrain. A framework for the evaluation of digital musical instruments. *Computer Music Journal*, 35(1), 2011.

J. Pressing. The micro- and macrostructural design of improvised music. *Music Perception: An Interdisciplinary Journal*, 5(2):pp. 133–172, 1987.

G. Schiemer and M. Havryliv. Pocket Gamelan: Interactive mobile music performance. In *Proc. IS-CHI '07*, pages 716–719. Research Publishing, 2007.

S. Smallwood, D. Trueman, P. R. Cook, and G. Wang. Composing for laptop orchestra. *Computer Music Journal*, 32(1):pp. 9–25, 2008.

H. Stenström. *Free Ensemble Improvisation*. Number 13 in ArtMonitor. Konstnärliga fakultetskansliet, University of Gothenburg, Sweden, 2009.

B. Swift. Chasing a feeling: Experience in computer supported jamming. In S. Holland, K. Wilkie, P. Mulholland, and A. Seago, editors, *Music and Human-Computer Interaction*, Springer Series on Cultural Computing, pages 85–99. Springer London, 2013.

B. Swift, A. Sorensen, M. Martin, and H. J. Gardner. Coding Livecoding. In *Proc. CHI '14*. ACM Press, 2014.

A. Tanaka. Mapping out instruments, affordances, and mobiles. In *Proc. NIME '10*, pages 88–93. University of Technology Sydney, June 2010.

D. Trueman. Why a laptop orchestra? *Organised Sound*, 12(2):171–179, August 2007.

C. Unander-Scharin, A. Unander-Scharin, and K. Höök. The vocal chorder: Empowering opera singers with a large interactive instrument. In *Proc. CHI '14*, pages 1001–1010. ACM Press, 2014.

G. Wang. Ocarina: Designing the iphone's magic flute. *Computer Music Journal*, 38(2):8–21, 2014.

G. Wang, G. Essl, and H. Penttinen. Do mobile phones dream of electric orchestras? In *Proc. ICMC '08*, August 2008a.

G. Wang, D. Trueman, S. Smallwood, and P. R. Cook. The laptop orchestra as classroom. *Comput. Music J.*, 32(1): 26–37, 2008b.

G. Wang, J. Oh, and T. Lieber. Designing for the iPad: Magic Fiddle. In *Proc. New Interfaces for Musical Expression 2011*, pages 197–202. University of Oslo, 2011.

D. A. Williams. Another perspective: The iPad is a real musical instrument. *Music Educators Journal*, 101(1): 93–98, 2014.

E. J. Williams. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 2(2):149–168, 1949.

J. O. Wobbrock, A. D. Wilson, and Y. Li. Gestures without libraries, toolkits or training: A $1 recognizer for user interface prototypes. In *Proc. UIST '07*, pages 159–168. ACM Press, 2007.

A. Xambó, E. Hornecker, P. Marshall, S. Jordà, C. Dobbyn, and R. Laney. Let's jam the Reactable: Peer learning during musical improvisation with a tabletop tangible interface. *ACM Trans. Comput.-Hum. Interact.*, 20: 36:1–36:34, 2013.

BFDLAC, A FAST LOSSLESS AUDIO COMPRESSION ALGORITHM FOR DRUM SOUNDS or SIZE ISN'T EVERYTHING

SKoT McDonald
FXpansion Audio (Australia)
skot@fxpansion.com
Perth, WA, Australia

ABSTRACT

BFDLAC is a fast-decoding, multi-channel, lossless audio compression algorithm developed by FXpansion for disk-streaming large drum sample libraries. It achieves near-FLAC file size reduction, but is over twice as fast to decode. These features are vital for use in real-time software instruments like FXpansion's *BFD*.

1. INTRODUCTION

BFDLAC is a new lossless audio compression algorithm developed by FXpansion for use in the 'BFD' acoustic drum software family. Our specific requirements were compressing multi-channel percussive sounds, disk-stream support, arithmetic simplicity allowing greater SIMD[1] CPU instruction use and parallel decoding.

Absolute priority is given to fast decode versus encoding times. We describe the BFDLAC algorithm, and compare its performance to other popular audio compression algorithms. BFDLAC is an algorithm that forgoes optimal performance in file size reduction and encoding time in return for a much improved decode performance.

1.1. Problem description

Sample-based software instruments have exploded in size over the past few decades. An Amiga "tracker" user in the early 1990s might have had their entire sound collection stored in a shoebox of 3.5" floppy disks, amounting to some tens of megabytes. With the ongoing "arms race" to provide ever-increasing detail and fidelity, it is now commonplace for a single software instrument to draw from hundreds of gigabytes. A typical musician may have many such behemoths installed.

The BFD family of drum software started with a 9GB core sound library in 2003. BFD2 hit the market with 55GB in 2007. In 2013, BFD3 launched with a 160GB wav-equivalent payload of sounds. A "drum kit" in BFD1 may have totaled 1GB; a BFD3 kit can span 40GB or more. BFD data is so large because every drum is exhaustively articulated: several playing styles, sometimes 100 or more velocity layers per style, and each hit recorded from 10-14 microphones. And then all that is repeated for different beater types – rods, mallets, brushes, etc.

Computer capabilities have improved in that time – most users' machines now run 64bit operating systems with 8 or more GB of RAM, and have fast solid state drives to deliver data. End-user performance expectations have also grown, possibly faster than their hardware; ever-more-sprawling expansion libraries appear, each demanding additional storage real estate. It is essential to find ways for a product to deliver more content of higher quality without massively increasing the proportion of a user's computing resources: storage, data bandwidth between storage and the CPU, and CPU power itself. Data compression can address the first two, but must do so without blowing out the CPU budget.

The product delivery model has also evolved with increasing connectivity and bandwidth. BFD1 was sold primarily through physical stores on 2 DVDs, uncompressed; BFD2 squeezed via data compression onto 5 DVDs. Nowadays, BFD3 is primarily sold as a direct download, despite being over 17 times larger than the first edition (when compared as uncompressed wav files). Digital downloads impose server and bandwidth costs on the manufacturer, as well as consuming customers' time and connection bandwidth. When dealing with payloads in the dozens of gigabytes going to thousands of customers, efforts to minimize payload size can quickly have real economic benefits. Compression is essential.

BFD Expansion Pack installers started using a customized version of the FLAC algorithm in 2005, with the launch of BFD Deluxe, a 55GB behemoth delivered on 5 DVDs. Unfortunately, FLAC decode was too slow on contemporary hardware for disk streaming, so BFD2 still required uncompressed WAVs. This might seem strange – FLAC decodes many times faster than real-time playback, but the sheer number of channels requiring simultaneous decoding in BFD can be taxing. A fairly uncomplicated drum pattern can cause 30-50 synthesis

[1] Single Instruction, Multiple Data instruction sets such as SSE and AVX allow modern CPUs to operate on several data simultaneously.

 ISSN 1448-7780

voices to play simultaneously, each with a dozen audio channels. BFD needs to be able to decode several hundred audio channels simultaneously. It must do so in a way that doesn't occupy every CPU core on a user's machine, particularly as musicians will insist on using several software instruments at once.

1.2. Solution requirements

- Get "close" to FLAC in file size reduction
- Be significantly faster than FLAC to decode.
- Disk-stream-able / random-access capable
- SIMD-instruction friendly
- Multi-core / parallel-processing friendly
- Support encoding for different target bit depths.

2. APPROACH

We are not trying to develop a general purpose audio encoder; rather observe and exploit features of the specific dataset in question.

Exponential decay: Single-hit drum samples are typically dominated by exponential decay envelopes. For cymbals and toms, a sample may extend for some tens of seconds. Sample producers additionally fade out samples over several seconds to minimise detection of an artificial end point. Large reductions are possible simply by variable encoding of bit depth.

"Simple" harmonics: Every sample is primarily a recording of a single resonating body being struck once. There are some mechanical and "through air" vibration coupling effects with other drums, usually at low levels.

Noise: Users have an ingrained perception that 24-bit audio equates to high quality. The human auditory system has a dynamic range of approximately 120dB, or 20 bits (Bregman 1994). It is rare for raw BFD recordings to achieve noise floors below -90dB (15 bits), due to the realities of sampling with analogue electronics. This is sometimes compounded by engineers recording through classic/vintage equipment that is also perceived as High Quality / Must Haves, by themselves or customers. Thus, we estimate that there is at best 8 bits of unavoidable noise, 4 bits of which are unperceivable to humans without dynamic range compressors. We may discreetly hold reservations whether The Customer Is Always Right, but they are the ones holding the money: hence vast effort is expended to encode and deliver electrical noise. Noise is definitely part of the natural realism any complete recording studio simulation must provide.

BFDLAC's encoding technique is built on the assumption that encoding and decoding the noise floor is where most effort should be spent. We break the signal into small blocks of 64 samples, and exhaustively search through a small set of simple approximation algorithm-plus-error-signal pairs, determining first which of these can actually fully encode the signal-block, and which results in the smallest encoded block.

The bit depth of the Error Signal is optimized per block, but is constant within the block, unlike variable bit rate Huffman-encoded Error Signals. All block encodings can be described as a pair of Sub-Algorithm and an Error Signal E of bit depth b, which result in a known, fixed block size for each pair.

This deterministic block structure and size allows for efficient SIMD assisted decoding, particularly of the error signal bit strings. For example, an Error Signal with bit depth of 6 stores 4 samples in 3 bytes; these 4 samples can be efficiently separated with SIMD bit shifting and masking operations applied in parallel.

3. FILE FORMAT

BFDLAC is a RIFF-like, chunk-based format derived from the Electronic Arts Interchange File Format, and described by Microsoft and IBM (Microsoft & IBM, 1991). An outer "BFDC" chunk contains a set of sub-chunks, as shown in Figure 1.

BFDC	fmt		
	Indx		
	data	Frame 0	Block 0
			Block 1
			…
			Block *n*
		Frame 1	Block 0
			Block 1
			…
			Block *n*
		…	…
	[custom]		

Figure 1. The chunk hierarchy of a BFDLAC file encoding *n* audio channels.

3.1. Format chunk ("fmt")

The format chunk encodes the following properties of the sound file:

- Channel Count (up to 255)
- Target Bit Depth (16,20,24)
- Sample Rate (44100 for BFD)
- Duration (in sample frames)
- BFDLAC version (currently 2, 1 is obsolete)

Format chunks are required, and additionally must be the first chunk in the file.

3.2. Index chunk ("Indx")

The index chunk provides a lookup table of data offsets at a specified interval of sample frames, with the aim of speeding up read access to locations within a sample without having to decode all proceeding data stream up to that location.

Index locations are encoded as a 32bit offset, at intervals of 2048 sample frames – 2048 being a common denominator of BFD's user-configurable disk-streaming buffer size.

Index chunks are optional, but should appear before the Data chunk if used.

3.3. Data chunk ("data")

The encoded data is composed of a string of *Frames.* These are not to be confused with sample frames, which represent a single simultaneous sample across multiple channels. Data chunks are required.

3.3.1. Frame

A *Frame* is a set of *Blocks*, with one block per audio channel. A Frame commences with a single byte encoding the number of audio channels. Since the count of audio channels doesn't change from the number given in the format chunk, this is also used as a simple check for data corruption or errors during decoding.

3.3.2. Block

A *Block* encodes the samples for a single audio channel within a Frame. In BFDLAC v2, block length is fixed to 64 samples. Each block commences with a header byte that describes which sub-algorithm is used (4 bits, 16 possible algorithms) and the bit-depth of an Error signal (4 bits, 16 possible error bit-depths).

All sub-algorithm/error signal pairs describe encoded data of a predetermined length. This allows easier parallel decoding of blocks; or skipping over the data of channels that are not being used, such as a muted channel.

3.4. Other chunks

Like all RIFF files, other optional custom chunks can be included to encode non-standard metadata such as BFD-proprietary meta data, authoring information, copyright notices and so on. BFDLAC encodes such information in the same manner as the RIFF-WAV format for familiarity. Optional chunks should be skipped if not recognized.

4. BLOCK ENCODING

4.1. Sub-encoding algorithms

BFDLAC is focused on fast decode, and is unapologetic about exhaustively testing all algorithm-error pairs to find the smallest encoding for every block, if needed. That said, many options are discarded quickly, and the encoding time remains competitive. We describe the 9 current algorithms restoring a signal y from starting coefficients x with a error signal E^b, b being bit-depth. Time, t, is the block-relative sample position.

4.1.1. Zero

$$y_t = 0 \tag{1}$$

Zero blocks are a special case to encode a block entirely of zeroes. It has no signal. Whilst it may seem like an extravagance to devote a sub-algorithm ID to zero, the frequency of which zeroed signal data needs encoding in BFD sounds makes it worthwhile.

4.1.2. Constant

$$y_t = x_0 \tag{2}$$

A constant, non-zero block. The value is encoded at the target bit-depth.

4.1.3. Median

$$\mathrm{y_t} = \tilde{x} + \mathrm{E_t^b} \tag{3}$$

A block with an error signal with median value $\tilde{x}$. The median is encoded at the target bit depth at the start of the block.

4.1.4. Quad

$$\begin{aligned} y_t &= x_t & t &= 4n \\ y_t &= y_{4n} + E_t^b & 4n+1 &\le t \le 4n+3 \end{aligned} \tag{4}$$

Blocks of length N are subdivided into n micro-blocks of 4 samples, with the micro-block initial value followed by 3 error values array of length $3N/4$ storing offsets. For a block size of 64 samples, there are 16 micro-blocks.

4.1.5. Oct

$$\begin{aligned} y_t &= x_t & t &= 8n \\ y_t &= y_{8n} + E_t^b & 8n+1 &\le t \le 8n+7 \end{aligned} \tag{5}$$

Similar to Quad, but with micro-blocks 8 samples long.

4.1.6. Diff

$$\begin{aligned} y_t &= x_t & t &= 0 \\ y_t &= y_{t-1} + E_t^b & t &> 0 \end{aligned} \tag{6}$$

A first order differential, with the value y_t equal to the previous value plus the error signal. A block is encoded with a starting value x_0 followed by $N-1$ error values.

4.1.7. Diff2

$$\begin{aligned} y_t &= x_t & t &= 0 \\ y_t &= y_{t-1} + E_t^b & t &= 1 \\ y_t &= 2y_{t-1} - y_{t-2} + E_t^b & t &> 1 \end{aligned} \tag{7}$$

A second order differential, where the gradient between two prior samples is used to predict the next sample position, and then corrected by the local error value.

4.1.8. HalfDiff2

$$\begin{aligned} y_t &= x_t & t &= 0 \\ y_t &= y_{t-1} + E_t^b & t &= 1 \\ y_t &= (3y_{t-1} - y_{t-2})/2 + E_t^b & t &> 1 \end{aligned} \tag{8}$$

A second order differential, with a scaling factor of 0.5 applied to the prediction gradient.

4.1.9. Gradient

$$y_t = A\,x_t + B + E_t^b \tag{9}$$

Linear regression is used to identify the gradient A and a constant B that best fit the data.

4.1.10. Uncompressed

If none of the other algorithms are able to encode the data, samples are simply stored at the target bit rate. Due to the block header byte, this results in 0.521% inflation for 24-bit samples with a block size of 64. Uncompressed blocks typically occur during the high-dynamic range, “noisy” onsets of sounds.

4.2. Block algorithm comparison

Table 1 shows the byte usage of the sub-algorithms for encoding signal coefficients and an 8-bit residual. Table 2 shows the frequency of which various sub-algorithms are utilized when encoding BFD3’s audio data. Note that commonly used algorithms Median, Diff and Diff2 have a very high percentage of their encoding bytes used for the Error signal. A competing algorithm that was able to encode the block with only 7 bit error could lead to a large saving, provided the signal coefficient overhead didn’t increase more than the saved error bytes.

Sub-algorithm	Signal	Error	Err %	Total
Zero	0	0	0.0	1
Const	3	0	0.0	4
Median	3	63	94.0	67
Quad	48	48	49.5	97
Oct	24	56	69.1	81
Diff	3	63	94.0	67
Diff2	6	62	89.8	69
HDiff2	6	62	89.8	69
Gradient	6	64	91.4	70
Uncompressed	192	0	0.0	193

Table 1. E^8 sub-algorithm byte utilization

Sub-algorithm	%	Sub-algorithm	%
Zero	35.63	Diff	33.25
Const	0.03	Diff2	10.13
Median	13.32	HDiff2	1.36
Quad	0.03	Gradient	4.69
Oct	0.01	Uncompressed	1.55

Table 1. Total incidence of each sub-algorithm at all error-signal bit-depths when encoding the entire BFD3 audio library.

Bit Depth	%	Bit Depth	%
Zero / Const	35.66	9	11.10
1	0.35	10	6.97
2	0.56	11	2.91
3	0.98	12	1.86
4	1.67	13	1.53
5	2.94	14	1.30
6	5.33	15	1.09
7	10.15	16	0.87
8	13.18	Uncompressed	1.55

Table 2. Total incidence of each Error-Signal bit depth across all algorithms.

5. ALGORITHM PERFORMANCE COMPARISON

5.1. Test

We wish to compare the encoding and decoding times, and resulting data payload sizes, of different compression algorithms. Encoding trial times are I/O inclusive, meaning start and end times include disk read and write access. Not all algorithms allowed us to time the data encoding entirely in memory. Decoding times are I/O exclusive, meaning we timed transcoding of a block of compressed data in memory to a block of uncompressed data, also in memory. We measured the decode times of only those algorithms that allowed such operations, and exhibited a useful amount of data compression.

5.1.1. Data

The test data was BFD3’s “Kit 1” dataset, comprising 37GB of 14 channel 24-bit wav files. Each sound file is a multi-microphone recording of a single hit of a drum or cymbal. Drums are hit in several styles, or articulations, and each articulation will contain 30-100 dynamics, or velocity layers. There are some 3206 samples in total, 5.5 hours of 14 channel audio, equivalent to over 3 days of single channel audio.

5.1.2. Machine

All data compression / decompression timing was performed on a machine with this specification:

CPU: Intel Core i7-4770K; 4 cores, 3.50GHz
RAM: 32 GB 1600MHz DDR3 RAM
OS: Windows 10
HDD: Western Digital “Green” 2TB 7200rpm SATA3

5.2. Comparison Algorithms

The following algorithms were selected for comparison:

5.2.1. Zip (Gailly and Alder 2013)

Zip is a popular implementation of the generic data compression algorithm LZ77 (Ziv and Lempel, 1977) and Huffman encoding. It is not specialized for audio; we include it as a baseline comparison.

5.2.2. RAR, RAR5 (Roshal 2015)

RAR is a proprietary generic data compression algorithm based on LZ77 and prediction by partial matching. It has optional “audio awareness”; a user can specify input data is comprised of a number of interleaved channels, which can aid encoding.

5.2.3. Tau True Audio (TTA) (Djourik and Zhilin 2004)

A block-based approach with inter-channel decorrelation, optimal selection of polynomial / linear predictive / adaptive filter signal modelling, and entropy (error) encoding.

5.2.4. Shorten (Robinson 1994)

Shorten uses blocked linear predictive coding with Huffman encoded error signal. Robinson suggests blocks of 256 samples for a 16kHz signal (6ms), compared to BFDLAC's 64 sample blocks for 44.1kHz (1.5ms).

Shorten is of interest as it uses a similar small-block approach with a variety of sub-algorithms, but unfortunately the current implementation of *Shorten* doesn't support 24bit wav files. *Shorten*'s 16bit encoding achieved similar compression ratios to *BFDLAC* 16bit.

5.2.5. Monkey's Audio (MA) (Ashland 2015)

Monkey's Audio converts stereo signals to a Mid-Side encoding, after which predictive coding is applied. The resulting signal and entropy is Rice encoded.

5.2.6. FLAC (Xiph 2005)

FLAC (Free Lossless Audio Compression) is a popular and patent-free codec using linear-prediction with Golomb-Rice encoding of any residual signal. It is our benchmark for best compression by absolute size. FLAC is also well-known as an easy-to-stream format. A custom modified version of FLAC is used by FXpansion to encode BFD audio data for installer payloads. FLAC is used because reducing the absolute size of data is paramount for economic data delivery. The modification we made allowed FLAC to support the large number of audio channels needed by BFD; FLAC is natively limited to 8 channels. BFD data payloads are transcoded during installation to WAV for older versions of BFD, or to BFDLAC for more resource-efficient use within BFD3.

5.2.7. WavPack (Byrant 2015)

WavPack uses an adaptive linear predictive coding model operating on stereo decorrelation, and a custom bit encoder for final compression.

5.2.8. BFDLAC v2 (McDonald 2015)

A block-based, sub-algorithm selective, bit-depth reduction focused algorithm that trades some compression size for greater speed and simplicity. Version 2, which appeared in BFD v3.1, has a fixed block size of 64 samples, 16 possible sub-algorithms (9 currently used), and supports error components encoded as simple PCM signals to all bit depths between 1 and 16.

6. RESULTS

6.1. Encoding BFD drum data

Table 3 shows the encoding time and resulting total file size of different lossless compression algorithms encoding BFD3's "Kit 1". Figure 2 shows the encoding time versus file size trade-off of the same algorithms. Table 4 shows the effectiveness of BFDLAC on the recordings of different types of drum. Note how effective BFDLAC is on "noisy" cymbals due to bit-depth reduction during very long decay tails.

Encoding	Time (s)	Size (GB)
Wav	-	37.975
Zip (best, 32kB)	927.38	16.041
RAR	1450.02	15.033
RAR5 (32MB cache)	1852.21	14.324
TTA	1746.00	9.460
Monkey's Audio (normal)	1482.03	8.872
BFDLAC v2	**522.67**	**9.510**
WavPack (extra)	1232.31	8.214
FLAC	910.07	8.020

Table 3. A comparison of encoding times and resulting data sizes of a variety of compression algorithms

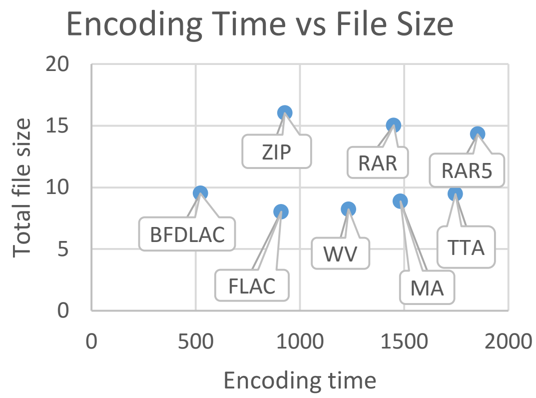


Figure 2. Encoding time versus compressed file size of various compression algorithms.

Class	Original GB	Encoded GB	%
Kick	1.767	0.515	29.10
Snare	5.267	1.492	28.33
Hihat	8.098	2.206	27.24
Tom	8.739	2.301	26.33
Cymbal	14.114	2.996	21.23

Table 4. BFDLAC encoding performance by Kit-Piece class of BFD3's Kit 1.

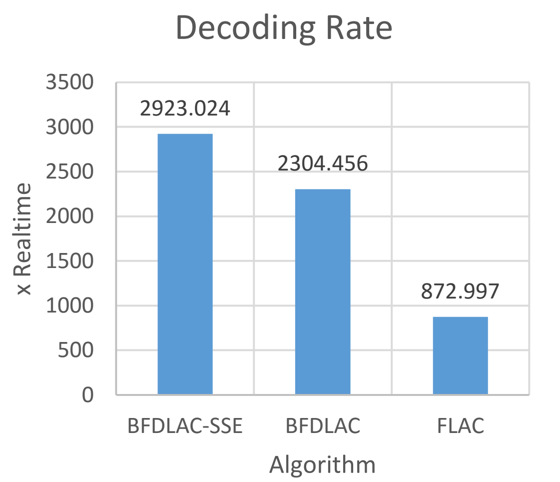


Figure 3. Comparing decode rate as a multiple of real-time playback of BFDLAC and FLAC encodings of BFD3's Kit 1.

Trade off (BFDLAC vs FLAC)

Algorithm	Disk space	Time	Real-time
Wav	100.0%	-	-
BFDLAC-SSE	25.25%	101.1	x2923
BFDLAC	25.25%	128.2	x2304
FLAC	21.12%	228.4	x873

Table 5. Resource use during playback BFDLAC versus Wav and FLAC. Time is in seconds (lower is better), Real-time is the number of channels one CPU core can decode simultaneously (higher is better).

6.1.1. Generic audio – "popular" music

BFDLAC is not designed to handle complex, mixed audio. Harmonically rich content defeats its limited ability to encode tonal sounds; its bit reduction features, which efficiently encode the long tails of percussive sounds, are little use on music with variable dynamics.

Song	WAV	FLAC	BFDLAC
Stairway To Heaven	135.6	62.1%	83.1%
Girl From Ipanema	80.9	35.4%	93.2%
Gangnam Style	56.6	49.2%	98.6%

Table 6. Comparing the encoding performance of FLAC and BFDLAC on "popular" music. Figures given are the final compressed file size as a percentage of the original wav file. WAV size is shown in MB.

7. CONCLUSION

We have introduced BFDLAC and shown it to be audio data compression algorithm that, while data-specific, yields a very beneficial compromise between optimizing for file size reduction and the time to decode data. In particular, BFDLAC-SSE trades a 4.13% reduction in wav compression performance for a 235% improvement in decode time versus FLAC.

BFDLAC is a core technology inside BFD, and crucial to BFD's delivery of industry-leading hyper-detailed and realistic drum sounds. BFDLAC enables finer-grained dynamics as well as additional drum articulations: thus more of the organic whole of an acoustic drum. On the engineering and production side, a larger palette of microphone channels gives engineers additional options for creative sound design. All this is achieved whilst reducing the computing resources needed, which of particular benefit for live performers such as e-Drummers using laptops.

Further performance improvements in both compression and speed are being actively realised, and more are anticipated. Correlation of stereo microphone pairs remains to be exploited; the work of converting more sub-algorithms and error signal decoding to full SIMD use is ongoing. Additionally, there is still scope in the file structure for deploying new block encoding sub-algorithms that result in reduced error signal bit-depth, and hence file size reduction.

8. REFERENCES

Ashland, M. 2015. *Monkey's Audio*. [Software]. Version 4.16. Available online at URL www.monkeysaudio.com. Accessed October 2015.

Bregman, A. 1994. *Auditory Scene Analysis*. MIT Press

Roshal, A. 2015. *WinRAR*. [Software]. Version 5.21, 64bit. Available online at URL www.win-rar.com. Accessed October 2015. win.rar GmbH, Berlin, Germany

Bryant, D. 2015. *WavPack*. [Software]. Version 4.75.0. Available online at URL www.wavpack.com. Accessed October 2015.

Djourik, A. and Zhilin, P. 2004. *Tau Producer.* [Software]. Available online at URL en.true-audio.com/. Accessed October 2015.

Gailly, J., and Adler, M. 2013. *ZLib*. [Software]. Version 1.2.8. Available online at URL www.zlib.net. Accessed October 2015.

Microsoft and IBM. 1991. *Multimedia Programming Interface and Data Specifications 1.0*. Available online at URL www-mmsp.ece.mcgill.ca/documents/ AudioFormats/WAVE/Docs/riffmci.pdf. Accessed October 2015.

McDonald, S. *et al*. 2015. *BFD3*. [Software]. Version 3.1.1.5. Available online at URL www.fxpansion.com/bfd3. Accessed October 2015. FXpansion Audio UK Ltd, London, UK.

Robinson, T. 1994. "SHORTEN: Simple lossless and near-lossless waveform compression". *Technical report CUED/F INFENG/TR.156*, Cambridge University

Xiph Org. 2005. *FLAC*. [Software]. Available online at URL xiph.org/flac/index.html. Accessed October 2015. Xiph Foundation

Ziv J., Lempel A. 1977. "A Universal Algorithm for Sequential Data Compression". *IEEE Transactions on Information Theory*, Vol. 23, No. 3, pp. 337-343.

GLACIER MELT: EMBODIED KNOWLEDGE AND THE COMPOSITION OF CHOREOGRAPHIC SOUNDSCAPES

Mark Pedersen *Prue Lang*
Victorian College of the Arts,
University of Melbourne

ABSTRACT

This paper presents an exploration on the role of embodied knowledge in composition process of soundscapes for choreography, utilising both subjective reflection and analytical techniques. Using Glacier Melt, a soundscape by Mark Pedersen for a choreographic work-in-progress by Prue Lang, the authors examine their own processes for aligning sound and movement, and compare these reflections to the subjective responses of dancers and audience members, as well as objective measures of the sound and movement material of the work in order to propose ways of cultivating greater awareness of and access to embodied knowledge in the composition process.

1. INTRODUCTION

The making of new artistic work always has a period of development, which frequently includes specific experiments and research into new techniques and processes. It was in the context of such a research and development phase that the authors collaborated, culminating in a limited public showing of a work-in-progress by Prue Lang that is part of her larger *Five Senses* project. The research questions Lang is posing within this development phase is to what degree can choreography be used to deliberately explore or provoke cross-sensory experiences, and what is the impact of such cross-sensory experiences on the reception of the choreographic work?

From a sound design perspective, the research questions of interest centre on the development of techniques for designing an aural environment for choreographic performance which are congruent to the affective goals of the choreography, and more specifically in developing a deeper understanding of the process of intermodal art making.

Frequently in sound design for choreography, there is a stark contrast between the processes used for art making. Dance making is, by nature, physical and much of the communication between the choreographer and the dancers is firstly via observation, demonstration and adjustment of physical moves, positions, and postures, secondarily underpinned by mental images and processes which are verbally communicated. In contrast, the sound design process is typically mediated through a computer interface and related interface tools, either in a studio environment or the rehearsal space, and as such the sound designer is frequently stationary, and at a distance from the dancers and choreographer. Communication is more frequently verbal or written, and furthermore, constrained by the lack of common terminology. At times, rather than sound designer and choreographer being able to work coherently toward a common goal, the sound design process is akin to stumbling in the dark hoping to find the light switch.

Embodied knowledge can be understood as being the type of knowledge in which the body knows how to act (Tanaka 2011), and is distinct from instinct or reflex, in that it typically arises from physical repetition (practice) and may sometime be referred to as muscle memory. Given the central role of embodied knowledge in the creation of choreographic work, the question then arises as to whether the tools and techniques of sound design can accommodate the transmission and use of embodied knowledge to overcome some of the communicative barriers in working across the different modes of expression within the two art forms. This paper combines subjective reflection by the authors with dancer and audience responses along with material analysis of the performance of the work-in-progress in order to explore this question.

2. CHOREOGRAPHY AND SOUND DESIGN

In the broader framework of contemporary choreography, sound design has taken on a significantly different role to that of the traditional musical score in ballet. Most famously, John Cage and Merce Cunnigham deliberately disconnected choreography and composition in their collaborative work. Likewise in the choreography of William Forsythe, ``music and sound are no longer relegated to the background of a dance but instead run independently in parallel to the organizing of bodies'' (Slater 2011).

In the work under discussion, the objective of the sound design is to support the choreography without distracting from it or competing with it. Which is not to say that the sound is always congruent to movement. Lang has a clear process of developing the choreography first, in isolation from the final sound design, and then introducing the sound design late in the development process. Frequently what is sought is sound design which offers a counterpoint to the movement, while still providing points of connection.

3. THE MAKING OF ….

3.1. Choreographic Context

Glacier Melt was composed as a soundscape for Part 1 of Five Senses, a five minute improvisation for four dancers, which featured slowly evolving responses to contact, at both an intra- and interactive level (i.e. contact between one body part and other and contact between dancers). Conceptually, one body part/dancer initiates contact and the other part/dancer receives information from the contact, but responses follow unexpected pathways.

The overall style of movement in this piece could be characterised as sustained, free flowing and indirect, with a tendency toward a feeling of heaviness. This mood is punctuated by moments of quick, light, direct movement that swiftly return to slow flow.

The choreography was well developed before the final sound design commenced, and in fact had not been seen by the sound designer before the composition of Glacier Melt.

3.2. Sound Design Process

The starting point for the design of Glacier Melt was a textual description of the choreography:

> “A soft very slow & fluid part with the hand sliding duets”

along with some suggestions for the sound quality:

> “something soft and meditative, circular … something that enhances/compliments”

Although the descriptive content of the text is quite sparse, this was sufficient to bring to mind a very clear sonic image for the piece. The sound design was completed in approximately 30 minutes based on two takes of improvisation with a combination of audio software (Native Instruments Reaktor) and modular outboard effects (Mutable Instruments Clouds). The resulting piece consists primarily of slowly cascading sine tones with occasional instances of harmonic distortion and delay.

Reflecting on the composition process, the role of embodied knowledge was significant. The brief textual description of “hand sliding duets” was familiar from the authors’ previous collaborations, and evoked a physical sensation, a mental image of the movement, and with those, a clear sense of what the piece should sound like.

While prior collaborative experience helped with the interpretation of the text, personal experience of somatic practices such as Tai Chi, yoga and Shintaido were much more dominant in shaping the internal awareness of the energetic and affective qualities that were sought in the sound design.

In this sense, the very sparseness of the textual description was helpful, in that it allowed room for Pedersen’s non-textual, embodied knowledge to inform the sound design process.

However, in examining the role of embodied knowledge in the sound design process, it is not particularly helpful to simply rely upon the happy alignment of personal experience and personal intuition of the sound designer with the intentions of the choreographer. The question we pose is whether embodied knowledge can be cultivated in support of sound design for choreography?

4. MOVEMENT, MATERIALISM AND MEANING

While the appeal to embodied knowledge, at first glance, appears to be a retreat into the purely subjective, there are also collective responses to sound that appear common and have their grounding in embodied cognition. The body is a starting point, since as embodied beings, human sound cognition is intimately linked with the way our brains also process movement, and as such, embodied sound cognition suggests that there are common elements to sound that underlie cultural and individual subjective layers of interpretation For in depth discussion of embodied music cognition, see Leman (2008) and Godøy (2010).

Three concepts are particularly useful to consider at this point:

- that sound is autonomous of audition, i.e. has its own materiality, independent of subjective experience
- that affect is likewise autonomous of subjective experience, allowing for individual, subjective experience while not denying the existence of a non-subjective reality
- that sound carries inherent meaning from the relationship it has to our own physical movement

This aspect of the relationship between sound, movement and embodied knowledge is recognised by Julian Henriques in *Sonic Bodies* (Henriques 2011), a detailed analysis of the systems of knowledge that underpins Jamaican sound system practice. Henriques proposes a model of knowledge which places the sensing, kinetic body at the centre of a sonic logos founded on ratio and relationship rather than representation.

Henriques’s concept of a sonic logos builds upon Lefebvre’s (2004) *rhythmanalysis*, embracing a broader framework of vibration as not just the means of propagating knowledge and affect, but taking the position that affect is vibration. Affect is carried across varying media, be they electromagnetic, corporeal, or sociocultural, and each of these layers of media manifests the rhythms, amplitudes and timbres of vibrational affect in different ways, be that certain patterns of audible frequencies, pulse rates, muscle contractions, patterns of movement, emotional responses, or cycles of social events, styles or fashions (Henriques 2010).

The same vibrational analysis is also found in Alfred North Whitehead's process philosophy, which underpins much recent significant theorising of sonic experience, from the work of Steve Goodman (Goodman 2009) to the work of Susanne Langer, as discussed by Priest (2013).

The vibrational-materialist approach places the locus of meaning within the relational web of vibrating and listening bodies. Affect finds its meaning in the relational, multi-sensory matrix of repeated patterns, arising from the repetition of a musical phrase, the enactment of ritual, the cycle of seasonal events.

Movement, sound and affect can be considered as different modes of the same vibrational relation - or rather the same vibration moving through different media.

This idea informed the composition of Glacier Melt in a practical sense, utilising prior investigation into qualities of motion and qualities of sound, particularly looking at Laban correlations of effort qualities and Smalley's spectromorphology (Pedersen and Alsop 2012).

Using these concepts as an investigative framework, we sought to look deeper into both the subjective responses and the material properties of the work in order to identify points of contact and transmission – conduits for the flow of affect.

4.1. Subjective Responses

Lang previewed the work at home while viewing a recording of a rehearsal. Glacier Melt was the closest of three offered selections, but she was unsure if it fitted the piece. Only when trialled with dancers in costume, lighting and music at the right level was there a sense of confirmation that the sound and movement matched the quality she sought for the work. The specific criterion for Lang in confirming the selection of Glacier Melt was that the sound design helped in delivering the right *live energy* within the space of the performance – the synergy between the choreography, the dancers, the sound and the space was key.

The dancers responded positively to the soundscape during final rehearsals and performance, with one commenting that the soundscape was ``complementary to the state [of being] the choreographer was looking to achieve… the idea and the music worked to create an atmosphere; personally this allowed me to find a better mind body connection: something to enter into as opposed to something [I had] to create''.

Audience responses to sound design for choreography are often sparse, and in as much as the intention for this piece was that sound should support but not dominate the choreography, this is to be expected and even welcomed. In the context of the performance under discussion, which was presented in a salon style setting in which the audience comments on the works at the end of the performance, audience responses were collected via informal verbal comments during the feedback session, and were followed by written email communication upon request.

Audience responses included noting that the performance and the sound both "felt cold" and "gave the impression of being under water". One member of the audience specifically noted the impact of the sound design, noting that they were struck by what they perceived to be "a perfect confluence of sound and movement … the constantly sliding pitch, neither opening nor settling in any particular tonality, and the way the movement flowed and slid and oozed in (apparent) response gave the impression that we, as an audience, were deposited directly into the middle of something that had been occurring in our absence...and would continue when we were gone again."

4.2. Objective Analysis

Figure 1 gives an overview of the sound and motion in terms of labelled sonic features overlaid on a spectrogram, with the amplitude wave form is shown below and a motiongram of the dancers' performance is shown above.

The motiongram (Jensenius 2006) was derived from a video recording of the performance, and shows the average percentage of motion within each 1 pixel wide horizontal band of the original video. Darker areas of density represent greater amounts of movement. Within the overall view in Figure 1, you can see changes in the amount of motion at different heights over time, with greater density of motion being at around torso height for most of the performance, and some obvious periods of greater motion low down (closer to the floor) just prior to the 2 minute mark and again just after the three minute mark.

Figures 2 & 3 compare more detailed excerpts from both the spectrogram and motiongram analyses. While there are no direct correlations between sound events and movement events, the qualities of each excerpt show some similarities, with the earlier material in Figure 2 being energetically calmer than that in Figure 3. Note the somewhat more intense harmonic strands and the presence of frequently upwardly rising spectral sweeps in the spectrograph component of Figure 3, indicating a significantly more complex soundscape at this point than in Figure 2. Likewise note the greater density of movement in the motiongram, including more significant vertical shifts in Figure 3 compared to Figure 2.

By way of contrast, Figure 4 shows an excerpt of a motiongram from Part 2 of Five Sense, featuring a more typical style of contemporary dance solo with many abrupt transitions and articulations. The motiongram reveals the much more dynamic style of performance in Part 2 compared to Part 1, with sudden gestures producing spikes and swoops in the motiongram. There was no musical accompaniment for this part of the choreography.

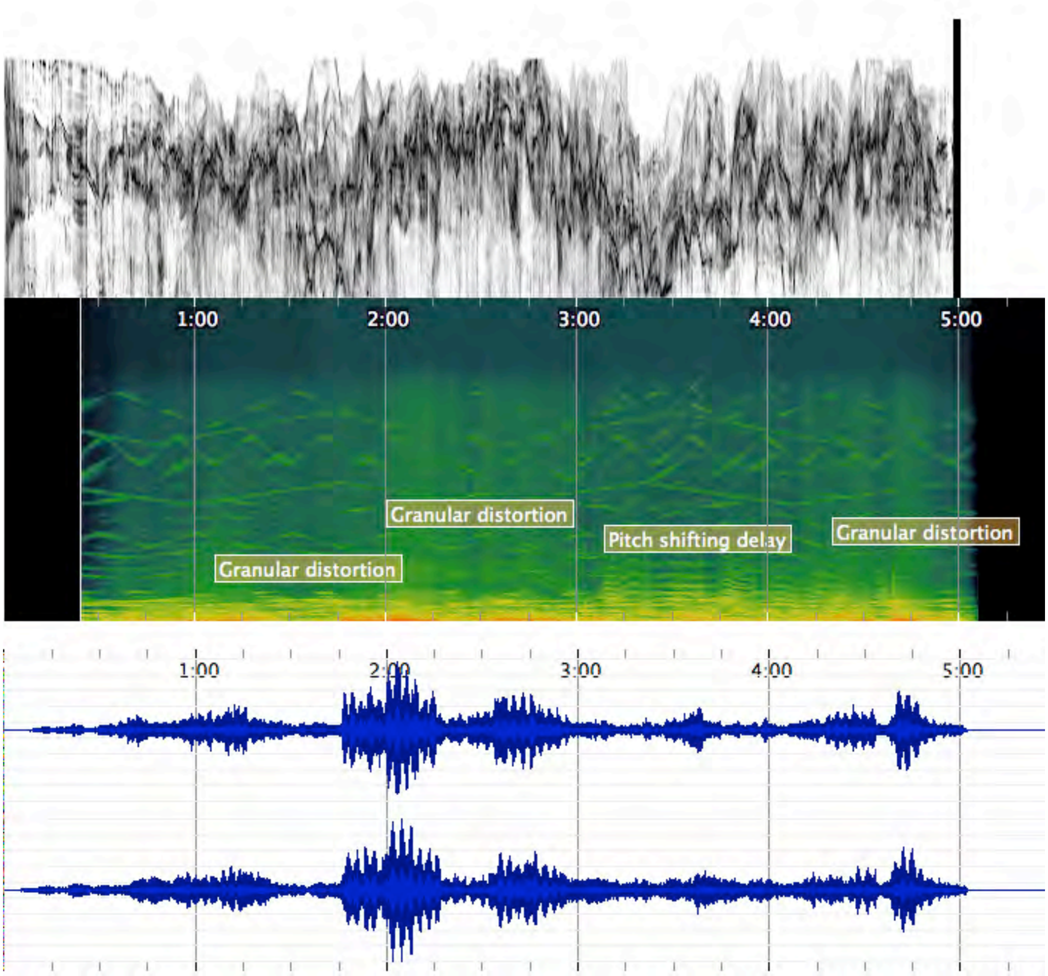


Figure 1. Spectrogram analysis of Glacier Melt aligned with motiongram of Five Senses (Part 1).

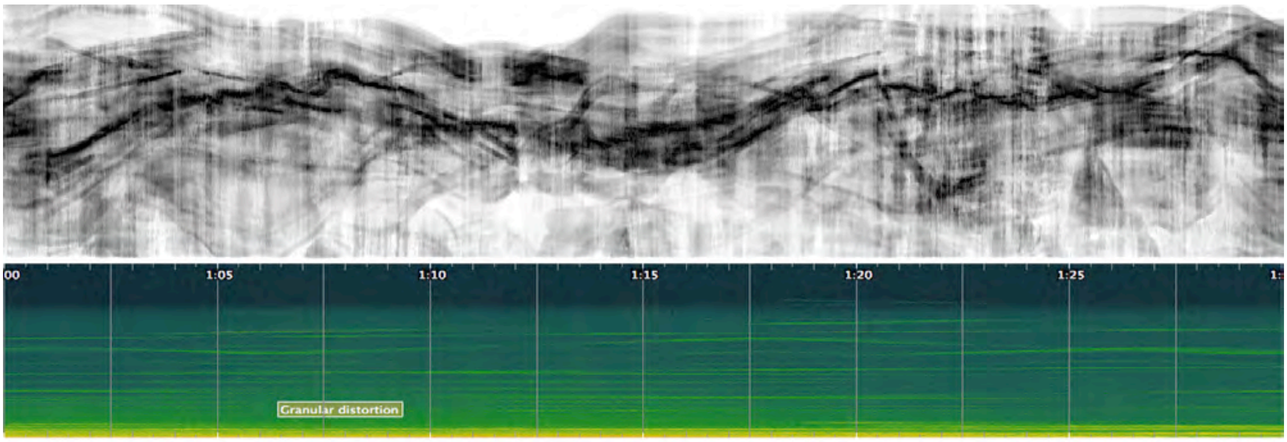


Figure 2. Spectrogram-motiongram excerpt 1 (1:00 – 1:30).

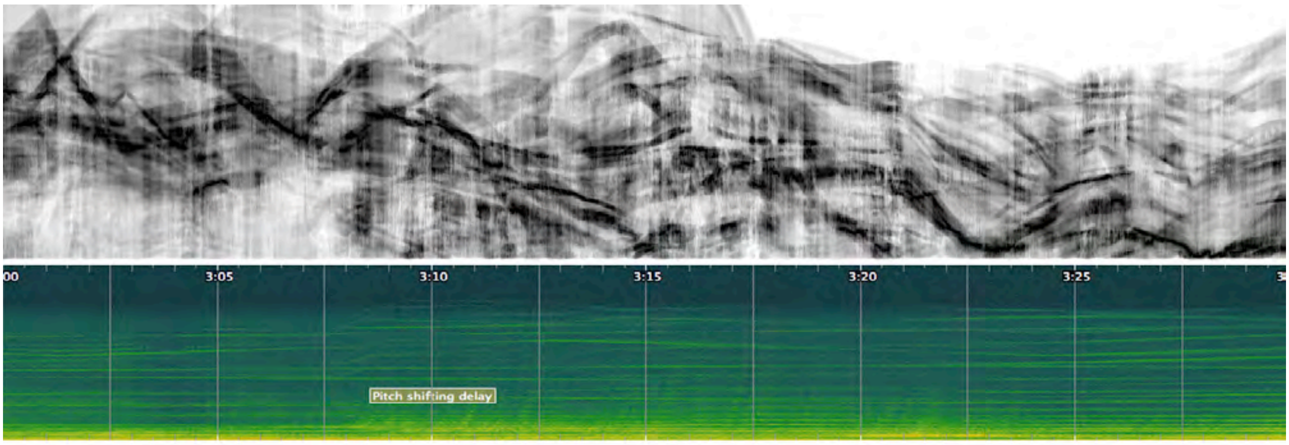


Figure 3. Spectrogram-motiongram excerpt 2 (3:00 – 3:30)

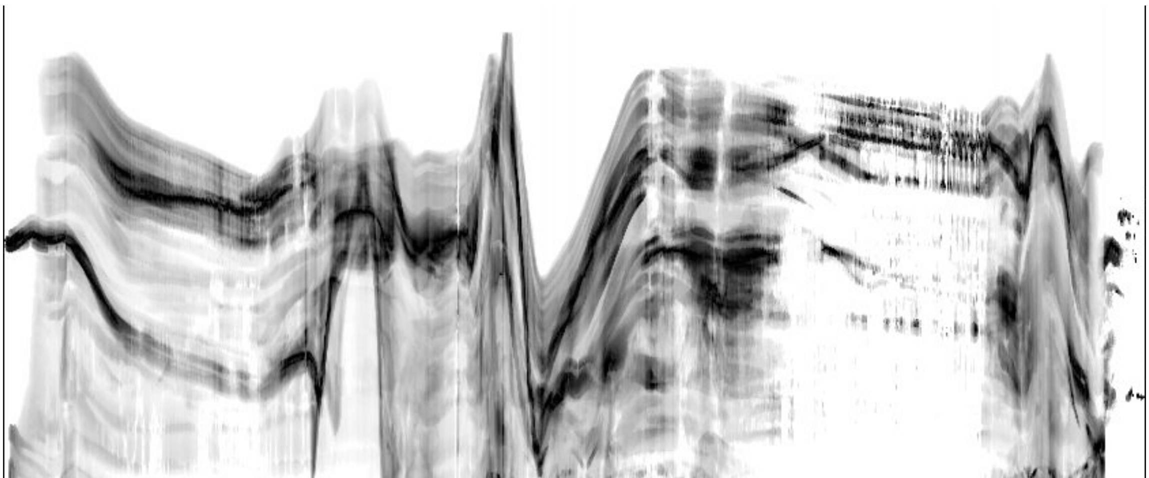


Figure 4. Motiongram excerpt from Part 2 (0:40 – 1:20)

4.3. Extrapolating from the abstract

Given the challenges in communicating verbally across modalities, the degree to which artistic collaborators share physical experience of each other's art forms helps to overcome the paucity of verbal communication. However, relying upon the embodied knowledge of dance that a sound designer may bring with them is inherently limiting. As noted in Section 3.2, there was a happy alignment of somatic experience and the choreographic content in Part 1 of Five Senses that enabled the composition of Glacier Melt to be informed by a congruent internal physical sense of the affective intent. For Pedersen, this degree of embodied knowledge would have been lacking if the sound design had been required for Part 2.

It is here that objective observations and analysis of one modality can help inform the creation process in another. By abstracting the complex physicality of the dancers' movement through motiongram analysis, qualities of motion can be observed and reflected upon in isolation from the choreography itself. This can be particularly useful if there is a preference for not constructing the sound design as a direct response to the choreography, as may occur if composing with direct reference to a video recording of a performance.

The motiongram in Figure 4 provides an insight into the dynamic qualities of that specific choreography, which can inform the sound design in ways that otherwise might not be possible, on account of either a lack of common verbal or notational language for communicating the style of movement or the desired sound qualities, or a lack of somatic experience with that style of movement.

While this approach to cross-modal knowledge sharing can be taken literally, and indeed, some research has been conducted into directly sonifying motiongrams, taking the motiongram as kind of generated score for the sound design (Jensenius 2013), we propose that an alternative approach is to *re-embody* the knowledge communicated abstractly by the motiongram.

The challenge with any re-interpretation of abstract knowledge lies in the mapping from abstraction back into a concrete form. In sound design, this can take the form of mapping (abstract) gestures to specific synthesis parameters, or, as Jensenius does, mapping from one graphic representation into a spectrogram and resynthesizing the generated spectrogram directly. In both cases, there can be much difficulty in the selection of appropriate synthesis parameters and the design of appropriate translations functions.

Working on the assumption that human beings are excellent transducers of affect, and that this is in large measure a function of embodiment, the next step in our exploration of cross-modal art making is to experiment with the use of motiongrams as guide for gaining personal somatic experience that is congruent with choreography as part of the sound design process. The goal is to enable the creation of strongly congruent pieces like Glacier Melt for choreographic material which might otherwise be difficult for the sound designer to assimilate.

5. CONCLUSIONS

We consider this just the beginning of examining the question as to whether the tools and techniques of sound design can accommodate the transmission and use of embodied knowledge to overcome some of the communicative barriers in working across the different modes of expression within sound design and choreography. By examining the subjective congruence of Glacier Melt as a soundscape for a specific choreographic form, we've found through the comparison of spectrograms and motiongrams what appears to be a reasonable objective analysis which is supportive of, but does not necessarily explain this congruence. The goal is not to follow a reductive process of seeking directly translatable metrics from one artform to another, but rather to find useful ways of understanding the subtle qualities of choreography and sound design in order to support the creation of cross-modal artworks.

To this end, further investigation is needed to test how these objectives measures can be used to support sound design for unfamiliar modes of movement (i.e. where embodied knowledge may be not be present to support the sound designer's understanding of the movement or the affective intent). One approach to this would be to use motiongrams to extract gesture profiles which could be explored by the sound designer in the context of their own practice.

This line of investigation will focus on whether the general observation of similarity at the energetic level in the objective data is sufficient to support the goal of creating other cross-modal works which achieve the same degree of congruence found in Glacier Melt. In particular, we propose to investigate ways in which such information is used not in direct translation from one material to another, but rather used to cultivate embodied knowledge within the practitioners, using the body as an integral part of the creation process in both choreography and sound design.

6. REFERENCES

Godøy, R. I., Leman, M., 210. *Music, Gesture, and the Formation of Embodied Meaning*, Routlege, New York.

Goodman, S. 2009. *Sonic warfare,* MIT Press.

Henriques, J. 2010. 'The vibrations of affect and their propagation on a night out on Kingston's dancehall scene', *Body and Society* 16(1), 57–89.

Henriques, J. F 2011. *Sonic Bodies: Reggae sound systems, performance techniques and ways of knowing,* Continuum.

Jensenius, A. R., 2006 "Using motiongrams in the study of musical gestures." *Proceedings of the 2006 International Computer Music Conference*. New Orleans: International Computer Music Association: pp. 499-502.

Jensenius, A. R., Godøy, R. 2013. "Sonifying the shape of human body motion using motiongrams", *Empicial Musicology Review, 8(2), 73-83.*

Lefebvre, H. 2004. *Rhythmanalysis: Space, time and everyday life*, Continuum.

Leman, M., 2008. *Embodied music cognition and mediation technology*, The MIT Press.

Pedersen, M., Alsop, R. 2012. "An approach to feature extraction of human movement qualities and its application to sound synthesis," in *Interactive: Proceedings of the 2012 Australiasian Computer Music Conference*, Brisbane, Australia.

Priest, E. 2013. 'Felt as thought (or musical abstraction and the semblance of affect)', in M. Thompson & I. Biddle (eds), *Sound, Music, Affect: Theorizing Sonic Experience*, Bloomsbury Academic.

Slater, C. 2011. "Timbral architectures, aurality's force: sound and music", in S. Spier (ed), *William Forsythe and the practice of choreography: it starts from any point* Routledge.

Tanaka, S. 2011, "The Notion of Embodied Knowledge and its Range" in *Encyclopaideia: Journal of phenomenology and education, 37*, 47-66.

PERFORMATIVITY AND INTERACTIVITY: CONCEPTS FOR FRAMING THE PEDAGOGY OF MAKING

Ian Stevenson
School of Humanities and
Communication Arts
Western Sydney University

ABSTRACT

This paper describes the production of a space for making within the context of undergraduate music training. It outlines attempts to develop an integrated conceptual and pedagogical framework for making electronic music. It maps practical responses to the problems at the centre of the development and delivery of a second-year undergraduate subject in music technology entitled *Machine Musicianship*. Its central conceit is that concepts and problems can be productive and stimulate creativity.

1. BACKGROUND

Tertiary music technology education is characterised by a great diversity of pedagogical methods, from theory-first approaches, to creativity and composition-led; from the rapidly disappearing one-on-one, to collaborative, group-based and student-centred models (Brown and Nelson 2014). As an arts practice discipline, music has also been caught up in sector-wide trends towards reduced funding and standards-based assessment practices (Tregear 2014; Morgan 2004).

This paper describes the production of a space for making within the context of undergraduate music training. It outlines attempts to develop an integrated conceptual and pedagogical framework for making electronic music. It maps practical responses to the problems that have arisen during the development and delivery of a second-year undergraduate subject in music technology entitled *Machine Musicianship*.

The problems that define such an undertaking include but are not limited to the following: how to orient students towards an experimental approach to music making where the lines between performer, composer and instrument builder are blurred; how to assess diverse musical outcomes and approaches where the results may not make for pleasant listening; how to encourage risk taking but value skills exploited in presenting work of quality; how to encourage concept driven work that develops the artist's independent voice; how to value contextualisation within broader historical, contemporary, global and local cultural practices; how to express a set of explicit and inclusive musical values that avoid invoking the musical prejudices of assessors and that are not opaque to students.

A response, in part, to these questions in the context of the *Machine Musicianship* subject is encapsulated in the problematic of performativity and interactivity, a pair or series of pre-paradigmatic concepts that are explored, developed and evaluated in the design of the subject's content, delivery and assessment. These problematic[1] concepts are taken up as assessment criteria that are explored in student-centred group learning activities that are intended to provide a demonstrable link between the learning outcomes for the subject, the content of the unit, in particular as stimulus for creative strategies and technical designs, and linked into an assessment framework.

2. A CURRICULUM FOR MAKING COMPUTER MUSIC

The subject under discussion is the fourth in a sub-major sequence in sound technologies delivered as part of the Bachelor of Music at Western Sydney University (formerly UWS). Over three successive semesters students are introduced to a range of potentially challenging repertoire from the twentieth and twenty-first centuries. These exemplars are intended to challenge their listening assumptions, open a discussion on the nature of contemporary music, introduce them to canonical works that enable them to engage in an informed way with the discourse of electronic music, and provide them with examples of how the practical topics in music technology that they are being introduced to relate to the practice of composers and performers.

The practical techniques covered in the first two semesters include conventional methods in stereo recording, MIDI sequencing, sound synthesis, and approaches to composition. These techniques are assessed in the context of practical projects framed by mainstream creative tasks such as performance documentation and production, sound design for sonic branding, and production music composition. These musical ideas are accessible to most musicians entering an undergraduate program. However, in what is probably a fairly standard approach to university-based music technology training, during the second semester students are stretched by being forced through assessment task descriptions to engage creatively with the everyday

[1] The term problematic is borrowed from Gilles Deleuze's reading of Kant in which problematic Ideas provide a "systematic unity" to an enquiry without expecting solutions or answers (Deleuze 1968/1994, pp. 168–169).

sound environment as a source for musical material that they must select, collect and with which they must compose.

The third semester offers an introduction to conventional multi-track studio techniques that allows students to develop skills that many identify as being important to them at the time of enrolment for higher education and that facilitates the development and presentation of their work in composition and performance subjects and enables them to produce a portfolio of recorded work that may be useful in establishing their professional careers. Interestingly, many students choose in this context to develop the more experimental approaches to the use of a broad range of sound material that they have been introduced to in their first year. These approaches could loosely be characterised as soundscape and acousmatic composition.

Throughout this process several forms of scaffolding (Sawyer 2006, p. 11) or staging (Collins 2006, p. 52) are taking place. Firstly, techniques, terminology and listening strategies advance hand-in-hand, each building on the other and enabling greater technical facility and access to effective methods of discourse and creativity. Secondly, the criteria that are offered to assist students to target their effort and by which students are assessed are introduced in stages in order to set standards and expectations that progress throughout the subjects. Criteria are aligned explicitly with assessment task descriptions and learning outcomes for the subjects. At first year level students are introduced to simple objective measures of audio production quality and the listening skills associated with their aural identification and description; the need for organised and systematic presentation of materials; academic referencing standards to encourage reading, listening and attribution; the development and expression of coherent conceptual frameworks for creative work; and a concern for musical form. These criteria and associated standards descriptions start with simple values and build and integrate as students progress through the curriculum.

This approach is not novel and has become best practice in higher education (Morgan 2004; O'Donovan, Price, and Rust 2004). The criterion and standards based assessment approach has been evaluated within the context of the curriculum described above (Blom, Stevenson, and Encarnacao 2015). As noted in that evaluation, many students use this scaffolding to target their efforts but many choose to ignore the institutional context of their learning. Many students aim just *to get through* while balancing the demands of earning an income and maintaining themselves and their families.

3. SUBVERTING THE ASSESSMENT FRAMEWORK

In an effort to achieve some of the pedagogical objectives outlined earlier the criterion and standards-based approach is somewhat subverted at the point that students reach the fourth semester. These pedagogical objectives include the production of a space for making within the context of undergraduate music training; orienting students towards an experimental approach to music making where the lines between performer, composer and instrument builder are blurred; encouraging risk taking; encouraging concept driven work that develops the artist's independent voice; and valuing contextualisation within broader historical, contemporary, global and local cultural practices. To achieve these objectives the learning environment is designed to develop what is known in the educational literature as *a community of practice* (Collins 2006, p. 51). This community shares a common set of goals including passing the subject and making music. Learning activities are structured so that students are thrown together to share and facilitate each other's learning. Social media is employed to stimulate an authentic sense of community and the normalisation of the learning experience[1].

The software tools employed in the subject present a problem and a challenge to this community. The media-programming environment Max/MSP (now simply Max 7) is alien, in some ways archaic and is difficult to learn. It harks back to a time when computer musicians had to build their own tools. It does encourage music making outside of the conventional musical paradigms. In truth however, new tools often appearing on tablet computing platforms are rapidly subverting these paradigms. In addition to allowing students to participate at some level in this subversion, the use of a patching environment is intended to develop some form of digital literacy (Jenkins 2009) without the need to write code. To this end it encourages systematic and programmatic logic, and a non-linear approach to music production.

A difficult and unaccommodating software environment is not the only problem that binds this community of practice. In addition to giving an introduction to music programming in the patcher environment, the subject aims to provide a conceptual understanding of interactive or responsive sound works, a practical understanding of performance interfaces for digital instruments and an ability to design, plan, realise and assess substantial creative projects. Problems proliferate and become a normalised part of work with music technology. A conceptual problematic defined by two key problems is embedded at the heart of the subject and expressed within the assessment rubric through which students must attempt to gauge their own progress and by which their assessors must evaluate their achievement. The two concepts of *performativity* and *interactivity* are explored throughout the subject and are tested for their ability to stimulate and produce new and interesting solutions in the form of creative works.

[1] In the last two years 98% of students enrolled are active users of facebook, whereas usage of the University's learning management system is limited. facebook usage has waxed and waned over the years.

4. WHAT IS PERFORMATIVITY?

The concept of performativity comes from the philosophy of language (Austin 1962, Searle 1979) and has since been adopted in a range of critical and cultural theory (Butler 1997). In the philosophy of language a performative utterance is one that brings a state of affairs into existence. This usually relies on an institutional context. Examples include the recitation of marriage vows and a proclamation by an authorized celebrant that produce a marriage in the act of speaking. Performativity links performance with creation, a concept not unknown the world of music, particularly in improvised performance. An improvisation is a unique instance of a musical work. Even in scored music a performance can be considered as an ontologically distinct token of a particular musical type that is brought into existence by its rendition in performance (Wollheim 1980; Wetzel 2011).

How then does performativity become a problem in electronic music performance? Electronic sound reproduction replaces the necessity for performance in the presentation of music. This challenge to the notion of electronic concert music was realized by the pioneers of electronic music in the middle of the twentieth century who sought a means to integrate some element of performance in the presentation of their music (Manning 2003). An aspect of musical authenticity is related to the production of sounds in real-time and this issue affects audience reaction to the performances of DJs as much as it does those engaged in experimental electronic music (Auslander 1999; Moore 2002, Emmerson 2007). One of the aims of new musical interfaces is to enable expressive and visually engaging musical performances (Paine 2015) that retain some of the embodied aspects of traditional instrumental performance.

Each compositional or performance strategy might be situated somewhere along a continuum of performativity where at one end is located the playback of fixed works and at the other are located works generated by real-time synthesis with continuous parametric control over the production of sound. An analysis of performativity in electronic music would likely reveal a great variety of approaches that would not sit easily on a single dimension. For example, real-time synthesis can be expressive and nuanced on a micro-scale, whereas interaction can produce variable structures on a larger scale. Highly engaging and successful musical works can easily combine more-or-less fixed sequences of material that are augmented by dynamic and expressive performance gestures on a limited range of vocal, instrumental or ornamental resources. There are no simple rules for performativity, however it is an easily understood value linked to the concept of authenticity and a useful way of problematizing electroacoustic music performance and stimulating creative design responses to the development of performance interfaces.

Performativity is expressed as an evaluative criterion with the subject documentation with the use of descriptive text associated with standards of performance. This text includes the following elements from poor performance to outstanding:

- Does not engage with the patcher environment as a performance or composition tool
- Mainly pre-recorded material and playback.
- Relies heavily on preproduction, or, comprises large blocks of pre-made material.
- Good balance between preproduction and real-time structuring on macro and micro scale.
- Most sound material or structures created in performance, strong link between performance gestures and sonic material.
- Highly nuanced technique and expressive performance.
- Unique realization created in the moment.

5. WHAT IS INTERACTIVITY?

Interactivity describes a situation where two elements within a system respond to each other in more-or-less predictable ways. Artist and academic Garth Paine (2002) surveys a number of models of interactivity relevant to experimental musical practice. He initially discounts merely responsive systems such as DVD player which responds to a button press. Citing Todd Winkler he introduces a three level model of interactivity:

- The conductor model in which a central musical intelligence commands a group of responsive performers coordinated by a score.
- The quartet model in which each player responds to the others, moment to moment with a form of organised but distributed musical intelligence that is coordinated by a score. Control can be subtly shifted from one player to another.
- The improvisation model in which the musical structure and be modified within an agreed framework and control is deliberately passed from one member to another.

Another model is that of the conversation in which each party responds to the less predictable contributions of the other with more-or-less spontaneous, novel and engaging results. As we can see interactivity and performativity are closely related concepts. They are both tied up with agency and novelty. Interactivity may be associated with stochastic algorithmic processes, with aspects of artificial intelligence, or with physical interfaces that present a wide range of control possibilities.

Interactivity is expressed as an evaluative criterion within the subject documentation with the use of descriptive text associated with its own standards of performance. This text includes the following elements from poor performance to outstanding:

- “Press to start” where no machine agency is present.

- Simple control systems inhibit expressive potential or variation in performance or limited machine agency.
- Adequate control and interaction or well developed algorithmic complexity.
- Effective interaction with performance system or sophisticated algorithmic material.
- Novel performance interface, interaction with sonic material or audience.
- Outstanding integration of acoustic and electronic elements.
- Advanced and expressive performance system.

6. CONCLUSIONS

The concepts of interactivity and performativity are debated by students and found to be unstable and problematic. They challenge students as composers and performers and stimulate creative responses to the challenge of developing novel musical responses in the context of exploring new software. Situating these concepts within the assessment framework opens the possibility or necessity of negotiation or demonstrating a claim for the validity of their own solutions to the assessment task requirements. This potentially shifts the authority from the assessor to the student and enhances their agency in their own learning and creativity. Of course many students prefer to ignore the institutional context and focus on their personal interests and needs, or find the notion of unresolved problems unsettling, but for some the challenge of new tools and new concepts is highly productive. This approach is intended to realize the ways in which concepts can produce a space for creativity and making.

7. REFERENCES

Austin, J. L. 1962. *How to Do Things with Words*. Oxford: Clarendon Press.

Blom, D., Stevenson, I., Encarnacao, J. 2015. “Assessing Music Performance Process and Outcome Through a Rubric: Ways and Means”. In D. Lebler, G. Carey, and S. D. Harrison, eds. *Assessment in Music Education: from Policy to Practice*. New York: Springer, pp. 125-139

Brown, A. and Nelson, J. 2014. "Digital Music and Media Creativities." In P. Burnard ed. *Developing Creativities in Higher Music Education*. Abingdon: Routledge.

Butler, J. 1997. *Excitable Speech: A Politics of the Performative*. New York: Routledge.

Collins, Allan. 2006. "Cognitive Apprenticeship." In K. Sawyer, ed. *The Cambridge Handbook of the Learning Sciences*. Cambridge, U.K.: Cambridge University Press, pp. 47-60

Deleuze, G. 1994. *Difference and Repetition*. Trans. P. Patton. London: Athlone Press.

Emmerson, S. 2007 *Living Electronic Music*. Farnham: Ashgate

Jenkins, H. 2009. "Confronting the Challenges of Participatory Culture: Media Education for the 21st Century." In *The John D. and Catherine T. MacArthur Foundation Reports on Digital Media and Learning*. Cambridge, MA: MIT Press.

Manning, P. 2003. *Electronic and Computer Music*. New York: Oxford University Press.

Moore, A. 2002. “Authenticity as Authentication.” *Popular Music*, 21(2): 209-223.

Morgan, C. 2004. *The student assessment handbook: new directions in traditional and online assessment*, New York: Routledge.

O'Donovan, B., Price, M. and Rust, C. 2004. “Know what I mean? Enhancing student understanding of assessment standards and criteria.” *Teaching in Higher Education,* 9(3): 325-335.

Paine, G. 2002. "Interactivity, Where to from Here?" *Organised Sound* 7(3): 295-304.

Paine, G. 2015. "Interaction as Material: The Techno-Somatic Dimension." *Organised Sound* 20(1): 82 - 89.

Sawyer, K. 2006. "Introduction: The New Science of Learning." In K. Sawyer, ed. *The Cambridge Handbook of the Learning Sciences*. Cambridge, U.K.: Cambridge University Press, pp. 1-16.

Searle, J. 1979. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge, U.K.: Cambridge University Press, 1979.

Tregear, P. 2014. “Enlightenment or Entitlement? Rethinking Tertiary Music Education.” In Platform Papers. Vol. 38, Sydney: Currency House.

Wetzel, L. 2014. "Types and Tokens." In E. N. Zalta, ed. *The Stanford Encyclopedia of Philosophy*. Available online at plato.stanford.edu/archives/spr2014/entries/types-tokens/. Accessed October 2015.

Wollheim, R. 19080 *Art and Its Objects*. Cambridge, U.K.: Cambridge University Press.

MAP-MAKING TOWARDS AN ONTO-CARTOGRAPHY OF THE DIGITAL AUDIO WORKSTATION

Michael Terren
Western Australian Academy of Performing Arts
m.terren@ecu.edu.au

Cat Hope
Western Australian Academy of Performing Arts
c.hope@ecu.edu.au

ABSTRACT

Levi Bryant's concept of "onto-cartography," (2014) or "map-making," presents a compelling framework for understanding complex systems, objects, and media. Map-making involves understanding these in terms of the operations of their constituent parts, what Bryant terms "machines." This "machine-oriented ontology" can be applied in order to map the digital audio workstation (DAW), a complex medium that pervades the creation of virtually all recorded music today, yet has evaded sustained critical investigation as a mediator of recorded sound. Using Bryant's four kinds of maps as a framework, the ways map-making might contribute to more nuanced understandings of the DAW, and how such a framework can lead to new compositional methodologies using the DAW, are explored.

1. INTRODUCTION

The digital audio workstation (DAW) mediates virtually all recorded music today. It's rapid rise to ubiquity in recorded music, especially electronic and popular music, demands new conceptual frameworks that provide better understandings around its role as a site of composition and as a mediator of cultural discourse.

Historically, until about the 1990s, the DAW was principally used for editing in recording studios, using expensive, purpose-built computers. Much of the processing and mixing was done using sophisticated outboard gear, relics of the analog era, and utilised by highly trained sound engineers and producers. It was presumed that the *composition* of the music took place prior to its insertion in the DAW (Pras, Guastavino, and Lavoie 2013). With the rise of the personal computer, much of the studio's functionality moved into the computer, and the DAW became a widespread site in which to create music.

The term "in the box" has been used to denote composition and production practices that entirely take place within the DAW. Many producers and composers today work entirely "in the box," without recording or mixing any sounds in an external, physical studio until the very final stages of composition, if at all (Power 2007). Despite the constraints implied by such a term, it can also point to a new way of composing—as Waters writes, "what musicians tend to be interested in and good at is using devices in a manner which operates at the edges of or outside the design brief" (Waters 2007). In other words, tools developed for the purpose of sound engineering and record production have been engaged for other means, such as composition.

As a popular site for the composition of recorded electronic music, the DAW is under-theorised. Barrett writes that all media are "encoded with historical knowledge and conventions and [are] therefore inextricably bound to conceptual and theoretical frameworks." (Barrett 2010, 191) If this is the case, then these frameworks are obscured by parties contingent to the DAW. A site of composition as popular as the DAW provides a new compositional methodology that deserves a greater interpretative nuance and a greater sense of care to its unique characteristics as a tool separate from its predecessor, the multitrack tape console. Composers, software developers, and human-computer interaction researchers could benefit from access to frameworks that can enable such productive interrogation of this relatively emergent medium.

This paper hopes to provide a sketch towards one such theoretical framework that interrogates the DAW, drawing from the critical and cultural theory of Levi Bryant, specifically his book *Onto-Cartography: An Ontology of Machines and Media* (2014). The framework proposed here is designed to interrogate the DAW from a posthuman, sociocultural perspective. The intention is to re-conceptualise the DAW as a compositional medium informed by many contemporary understandings of sound, music, software, and technology. In doing so, it can offer a blueprint through which composers can use the DAW in ways that interrogate its mediatic condition. A key benefit of this framework is that it doesn't prioritise conventional or even intended understandings of the DAW—this epistemological shift can enable new techniques and appropriations unique to the DAW. This positioning of the DAW also seeks to assuage the limits of the current conceptual resources in the fields of human-computer interaction and affordance theory, as would be more commonly deployed in researching new musical interfaces. These limitations will be addressed throughout.

Mark Fell writes that "any tool is subject to redefinition through its uses, and dependent on its placement within wider social and cultural contexts; for example, my Dad's use of a screwdriver to open a tin of paint, or a friend's use of a shoelace to commit suicide." (Fell 2013) Here, Fell touches on the *pluripotency* of tools, a fundamental component of Levi Bryant's "machine-oriented ontology," described below.

2. ONTO-CARTOGRAPHY AND MACHINE-ORIENTED ONTOLOGY

Object-oriented ontology (OOO), also described as speculative realism, has surged in the last decade to become a principle point of contention in contemporary philosophy and metaphysics. Spearheaded by Graham Harman, Quentin Meillassoux and Levi Bryant amongst others, OOO emerged in opposition to the discursive, deconstructionist philosophies of the likes of Jacques Derrida and Slavoj Žižek, to advance realism and materialism as concepts that can and should be taken seriously in contemporary thought. A significant feature here is the decentring of the human subject as the inevitable mediator of reality, a trend usually associated with post-humanism. OOO, while accepting the inability for humans to perceive "reality," argues that questions of reality and ontology shouldn't end there. Thus, the project of OOO is the speculation on material reality as a distinct possibility. Meillassoux, one of OOO's central proponents, argued that philosophy since Kant has been principally disposed towards placing the human subject at the centre of all questions on the nature of being. Meillassoux called this stance "correlationism," because it is assumes that objects can only be explained in correlation with the mind of the human subject (Meillassoux 2009, 5). Correlationism is an anthropocentric concept that creates what the object-oriented ontologists believe is an unnecessary divide between human and world, subject and object. This correlationism is prevalent in many fields of inquiry—in relation to human-computer interaction, Ian Bogost writes, "[human-computer interaction studies] concerns itself with *human*-computer relations, not computer-computer relations. … Despite its technical tenor, computing is just as correlationist a field as everything else, obsessed with human goals and experiences" (Bogost 2012, 107).

Levi Bryant suggests that a more important ontological question than "what are objects?" is "what can objects do?" This shift in focus informs the basis of his "machine-oriented ontology." His central proposition is that material reality is composed entirely of machines, a term explained below.

2.1 Machines

A machine is a "system of operations that perform transformations on inputs thereby producing outputs." (Bryant 2014, 38) Adapted from Deleuze and Guattari's use of the term, Bryant uses the term "machine" instead of "object" not only to distance the concept from anthropocentrism (as the term "object" tends to be coupled with a "subject"), but also to emphasise the dynamic becoming of all material processes. A tree is a machine that transforms water, soil nutrients and sunlight into material outputs such as leaves, wood, oxygen and homes for various organisms and animals. A tree can accept many different inputs and transform them in a number of ways—for example, tying a rope and a tyre to a branch on a tree could produce a child's swing as an output. A chainsaw could fell a tree and produce fuel for a home's fireplace.

Machines are composed of other machines, and can assemble to form greater machines, but the hierarchy implied here isn't necessarily neat and tidy, and machines rely on each other's outputs to greater or lesser degrees. Bryant calls this the "gravity" of machines, in which machines can exert a strong influence over other machines that may be disabling for the subordinate machine—for example, a tree simply cannot survive without water, thus water exercises tremendous gravity over the tree.

Machines are understood not only by what operations they *do* (their products), but what they *can* do (their "powers"). Bryant distinguishes these as a machine's "local manifestations" and "virtual proper being" respectively. In doing this, Bryant sets up a degree of independence between a machine's *powers* and the outputs it produces. To treat these separately as Bryant does immediately presents the possibility of redefinition of machines we know well. A screwdriver may well be designed to fasten materials together by means of a screw and a human's twisting arm, but any definition of the human-screwdriver-machine will need to take into account its capacity to open a tin of paint, or be used as a weapon, or mix orange juice and vodka together. The virtual proper being of a screwdriver need not be limited to its manifestations as they correlate to human subjecthood—the screwdriver may operate on water to produce rust, or the handle may operate on light to produce photons of a certain wavelength, or it may operate on a magnet to produce magnetic attraction. This is what is meant by the *pluripotency* of machines—they will always have more possible manifestations than their *intended* ones. This concept is related to that of the more frequently deployed concept of affordance, however affordance tends to carry anthropocentric connotations, in which objects are positioned as enabling or disabling possible uses for the human subject only, as is insinuated by Mooney when he defines affordances by what they allow "us" to do

(Mooney 2010). Pluripotency relinquishes the necessity of the object/subject divide. By understanding the powers and products of machines in this way, object-object relations are given exactly as much priority as object-human relations.

Machines are “structurally open and operationally closed,” (Bryant 2014, 54) in that they are open to any number of inputs or flows, but closed in the sense that a machine “always transforms that flow according to its own operations and ‘processes’ those flows in terms of the internal structure of the machine.” (Bryant 2014, 56) Consider the soundwave: while humans are able to cognitively perceive and comprehend sound within elaborate conceptual and sociocultural frameworks, a screwdriver lacks these faculties, thus it transforms the sound, shearing it of all its (human) signifiers, in order to operate on it in a purely vibratory, physical capacity. Elephants and bats, known for their infrasonic and supersonic hearing respectively, transform sound in ways different to the human and the screwdriver.

3. MAPPING THE DIGITAL AUDIO WORKSTATION-MACHINE

Bryant is primarily interested in deploying onto-cartography in the service of a “thermodynamic politics,” (Bryant 2014, 72–74) by attending to and deconstructing the constituent machines of undesirable ecologies, such as capitalism, heteropatriarchy, or global warming. This call-to-arms is brought about by what Bryant perceives to be the stymied state of contemporary critical theory, which is perfectly able to diagnose the problems of contemporary society, yet unable to specify effective remedies for them. (Bryant 2014, 274)

Yet this machine-oriented ontology framework can be applied specifically to the DAW. In an onto-cartographical framework, the DAW is an assemblage of machines in varying states of relation, mediation, and becoming. However, to describe the DAW as a self-contained system of relations would be insufficient. There are, after all, many machines that interact, and the DAW can and does interact with broader semiotic, sociocultural, aesthetic and political machines.

For Bryant, the onto-cartographer is concerned with creating maps with the aim of understanding the way machines relate and interact in order to “provide us with the means to constructively intervene in worlds so as to produce better ecologies or assemblages [of machines].” (Bryant 2014, 257) To this end, he devises four kinds of maps, each one mapping machines and their relations with respect to different temporal positions. These maps can not only diagnose problems within current machines, but offer a starting point for the devising of pragmatic methods for the deconstruction of “bad” worlds, and the creation—what Bryant calls *terraformation*—of “better” worlds. These maps can provide new ways for composers to intervene in the DAW in the creation of new work, and also provide software designers a means to re-evaluate the possibilities of the DAW.

3.1 Topographical maps

A topographical map provides “a sort of snapshot of worlds or assemblages at a particular point in time.” (Bryant 2014, 259) Creating a topographical map involves identifying four things: the machines that constitute an assemblage, their hierarchical relations, their sources of energy, and their material outputs.

Identifying machines of the DAW is not a particularly difficult task. One of the machines most characteristic of the DAW's operation is the "audio track," a container in which an audio file is placed to facilitate the playback of that sound. The audio file, represented by a waveform, can be moved on a horizontal axis by the user for the sound to play at different times relative to the starting point (0:00). The track spans the horizontal length of the screen—theoretically there's no end to the possible temporal expanse of the track. The track is mediated by the “channel,” a series of processes and controls—machines—that manipulates the contents of the track as they play. A channel’s volume control will lower the amplitude of all audio on the track, and a panorama control will move the audio around the stereo or surround field, in effect changing the amplitude of each individual speaker. A "mute" button can prevent the track’s audio from sounding at all. Each of these controls can themselves be controlled with respect to time by another feature of the audio track: automation. A coloured breakpoint function, drawn with the mouse by the user, can enable the volume, or panorama, or any number of other functions, to take on certain values at certain moments throughout the timeline. Many automations can cohabit the same track, in which case several horizontal "lanes" of automation underneath the audio track can represent different parameters.

Channels can also contain plugins that mediate the audio in a number of ways. These may effect the timbre of the audio, as in the case of the equaliser; or adjust the dynamic range of the audio, as in the compressor; or even replay the audio at a predetermined interval, as the delay effect does. The linear hierarchy of plugins on a track is strict, however on plugins such as compressors it can be bypassed by receiving audio from another track using what is known as a sidechain. Finally, all these concurrent tracks and channels are combined in a process often called “mixing,” into a single, “master channel.”

This may give an idea of the labyrinthine complexity and hierarchical relations that a topographical map of the DAW entails. While the relations described are mostly those *prescribed* by software designers—themselves derived from the design of the multitrack tape console—that isn't to say that they are the only possible relations. Another example of this complexity comes from Cycling '74's "Live Object Model," a diagram that details the hierarchies of most of Ableton Live's parameters, which enables Max programmers to utilise the LiveAPI framework as part of Max for Live. (Cycling '74 2015)

One onto-cartographical exercise that can be applied to the DAW for a variety of unexpected relations to emerge, has become something of a tradition in OOO-based literature: the list, also described as the litany.

3.1.1 The litany

OOO-based literature deploys the list as a way of bringing disparate entities together into a field of relations. Bogost calls these productive lists "Latour litanies," after Bruno Latour's extensive use of them, (Bogost 2012, 30) and on that basis he developed an applet called the Latour Litanizer, compiling a short list of titles of random Wikipedia pages, encouraging the user to work out how these entities could possibly relate (Bogost 2009).

Continuing in this tradition, a list of machines in the DAW can be proposed, inviting the envisioning of how such machines *could* relate, despite them not being designed to relate specifically and directly. The key is understanding that all machines are pluripotent—we cannot know everything machines *can* do.

DAW Litany: Piano roll, "insert silence", sidechain compression, tap tempo, automation, zoom in, solo track, audio waveform, panorama, freeze track, undo, transient detection, EQ, duplicate, record arm, master track, quantisation, send effect amount, pencil tool, Rewire, CPU usage display, buffer size, software synthesiser, volume fader.

A litany such as this discloses two facets about the DAW, and of complex systems in general. Firstly, the litany's random assembly can reveal unusual relations that could potentially become an impetus for new compositional methods. Finding or creating new relations between, say, the "tap tempo" and panorama controls invites a composer to experiment with how they can relate in practice. Secondly, it reminds us that the DAW consists of machines, and while these machines may be relating to one another in different ways, they remain distinct and specific—machines are never defined simply by their relations to other machines. (Bryant 2014, 181). "Lists remind us that no matter how fluidly a system may operate, its members nevertheless remain utterly isolated, mutual aliens." (Bogost 2012, 40)

3.2 Genetic maps

Genetic maps are another of Bryant's map types, that "chart the genesis or history of how particular worlds came to be." (Bryant 2014, 263) Histories of the DAW do exist, but they often reduce to a chronological listing of new technologies as they are released, incrementally more powerful than their predecessors. (Holmes 1985, 287–88) A genetic map of the DAW would, as Jonathan Sterne compellingly argues, need to take into account "the social and cultural conditions that gave rise to these sound [technologies] and, in turn, how these technologies crystallized and combined larger cultural currents." (Sterne 2003, 2) Such genetic maps of sound technologies are becoming more commonplace, thanks in part to Sterne's advancement of sound studies as a viable avenue of cultural inquiry (see Sterne 2012).

One particular historical example as it applies to the DAW is that of its predecessor, the multitrack tape console. As Duignan notes, the DAW's design takes many cues from the design of the multitrack tape consoles of pre-digital recording studios. (Duignan 2008, 51–74) Tracks are striated and ascribed their own independent parameters and hardware/screen real estate. The linearity of the track is readily apparent in the DAW—even terminologies developed for tape such as "splice" and "snip" are used as metaphors to describe software editing procedures. Metaphors, for personal computer software designers, were thought to enable this efficient, frictionless interaction between the master user and the slave-like tool. (Barr 2003) This emphasis on metaphors is what Bolter and Gromala called an "aesthetic of transparency," where programmers set out to design interfaces that would "serve as a transparent window, presenting the user with a workspace without interference or distortion." (Bolter and Gromala 2006, 375)

As Pold writes, software was, and still is, utopically envisioned to become "a more or less direct coupling on the consciousness allowing the machine the ability to establish a direct connection between mind and machine … [becoming] the perfect butler, partner or surveillance apparatus." (Pold 2011, 96)

Of course, such a transparency never existed, and if we accept Bryant's machine-oriented ontology, then such a "frictionless" interface is literally impossible between operationally closed machines. This design aesthetic heavily informed the DAW, according to Duignan, and any genetic map of the DAW should thus take software aesthetics into account. A composer may use such a genetic map of the DAW and its metaphoric relationship to analogue media in a way that antagonises or problematises these "analogue" tropes inherent in the DAW, thus potentially discovering new relations amongst machines in the DAW.

3.3 Vector maps

A vector map is essentially a futurological analysis, a map that "chart[s] the trajectories along which worlds are unfolding." (Bryant 2014, 265) Like all futurologies, there are varying degrees of fallibility. An example of a widely agreed-upon vector map is that of global warming in the 21st century. Due to anthropomorphic greenhouse gas emissions, if no intervention is enacted, Earth will see average temperatures rise by over 2°C by 2100. The implications of such a temperature rise—the melting of the polar ice caps, the destruction of reefs, rising sea levels, etcetera—provides the necessity to interfere in the present, in order to avoid such an adverse future. The stakes, of course, aren't quite so high in mapping a futurology of the DAW. In fact, such futurologies are rarely expounded in current art discourse. A simple example of a vector map might be to say that as professional "creative" software increasingly becomes more intertwined with internet or cloud functionality, the DAW will follow suit. A subscription-based model to use a DAW, similar to Adobe Creative Cloud's model, might have significant consequences for the mobility of a composer, who cannot open such software in locations that lack internet access.

A more complex vector map can be elucidated by turning to other theories of media.

3.3.1 Post-DAW music, post-DAW interfaces

Maras and Sutton explore the trajectories (or vectors) of media, and the fuzzy transition of media from "emergent" to "established." (Maras and Sutton 2000) According to Maras and Sutton, new media tend to undergo a period in which imitation of established media is done in order to legitimise the new medium as capable of expressiveness in much the same way. This tends to be followed by a period wherein work in this new medium engages in forms and structures unique to its physical characteristics. A common example is photography's relationship with painting—photography as an artistic medium usurped the painting's depictions of the real, before developing uniquely photographic tropes that play with the material reality of photography.

After the emerging media engages in this "medium specificity," the established media may lose fixity. Here, Maras and Sutton describe the movement towards the use of the prefix "post," "indicating a questioning of foundational materiality of the medium." (Maras and Sutton 2000, 105) Thus, painting develops impressionistic or abstract styles in reaction to photography's effortless depiction of "real" visuality. As the authors note, it's important to remember that media are perpetually in flux, and "no medium exists in a final form," (Maras and Sutton 2000, 104) thus mediatic relationships aren't so clear-cut.

In this understanding, it would take new music tools and compositional interfaces to emerge, innovate and re-imagine themselves in order for the DAW to reciprocate. We are starting to see this, with the re-emergence of electronic music being made outside the DAW, as can be seen in analog synthesis, modular synthesis and drum machines in electronic music recording and performance. This can be for a number of reasons—Stuhl writes that digital tools are still widely seen as unsatisfactory emulations of their more "authentic" analog counterparts, (Stuhl 2014) while practitioners such as Richard Devine have said that the DAW encouraged neurotic and un-productive over-editing, that the relative confines of modular synthesis didn't enable. (Future Music 2013)

Perhaps, then, these trends and viewpoints are setting the stage for the emergence of a post-DAW music—a music that derives new outputs from the DAW's virtual proper being. The specifics of what and how such a music will sound are beyond the scope of this paper, but the onto-cartographical framework presented here might be able to explain such material engagements and how they differ from music currently made with the DAW.

3.4 Modal maps

Whereas vector maps describe a world in the future without intervention, modal maps describe *possible other* worlds. Modal maps "map futures that *could* exist if we intervene in ecologies in particular ways." (Bryant 2014, 266)

The vector map works on a linear conception of time—a straight line from past to present, then extrapolating further towards some endpoint where the vector map resides. Modal maps, on the other hand, can be thought of as nonlinear in their possibilities stemming outward from the present. Modal maps operate on the premise that all events are contingent, that everything that has happened *could* have happened differently, unlike the relatively strict causality implied by the vector map.

A modal map can be described as a utopic futurology—a best possible outcome if intervention is enacted. Thus, it tends to be "the domain of activists, militants and generals." (Bryant 2014, 266) Creating such a map of more socially significant machines, such as the state, might seem more easily applicable here, but the question of what a best possible future of the DAW might be isn't an irrelevant question to ask. The DAW isn't an untarnished medium—as Bell points out, the prevalence of DAWs like GarageBand threaten to centralise music pedagogy, in which "software developers [become] the music educators with the greatest reach and influence in the computer-dependent world." (Bell 2015) Magnusson points out that the DAW and its popularity demands alternative interfaces and possibilities for electronic music, lest the DAW epitomise the "fossilization of music into stylistic boxes." (Magnusson 2010) Magnusson's modal map thus involves a plethora of new user interfaces that negate the linear, repetitive and segregated nature of the DAW and its track system.

Interactions between the DAW and sociality are of greater importance for the modal map than problems with the DAW's user interface (as explored in great detail in Duignan 2008). A modal map of the DAW should be cautious of the DAW's complicit role in these various social circumstances, assuring that inequity is not introduced or exacerbated.

4. CONCLUSION

Bryant has little to say about the role of artistic practice in onto-cartography. He writes that "a great work of art is plastic in the sense that it is *pluripotent*. It is a machine that is capable of resonating in a variety of ways given the historical and cultural milieus that it encounters." (Bryant 2014, 52) This emphasis on the pluripotency of art is appropriate, however the implication that art only becomes "great" through its perserverance through time, through different milieus, is conservative. What if art becomes "great" through its reappropriation of machines?

Bryant's four map framework provides a way to think about the DAW when it is engaged in composition. To what extent can the artwork be the *map*, and its creators its map-makers, rather than just machine operators? Of course, maps are themselves machines that can lead to material manifestations if used in their intended way, that is, the deconstruction and terraformation of machines. (Bryant 2014, 124)

The framework outlined in this paper posits four ways to plot out and understand machines and assemblages of machines, specifically with the intention of instantiating new discourses and compositional techniques, upon a medium that has evaded sustained critical inquiry considering its privileged position as the mediator of most recorded music today.

5. REFERENCES

Barr, P. 2003. "User-Interface Metaphors in Theory and Practice." Masters diss., Victoria University of Wellington.

Barrett, E. 2010. "Appendix." In *Practice as Research: Approaches to Creative Arts Inquiry*, edited by Estelle Barrett. I.B. Tauris.

Bell, A. P. 2015. "Can We Afford These Affordances? GarageBand and the Double-Edged Sword of the Digital Audio Workstation." *Action, Criticism and Theory for Music Education* 14 (1): 44–65.

Bogost, I. 2009. "Latour Litanizer." *Bogost.com*. December 16. http://bogost.com/blog/latour_litanizer/.

Bogost, I. 2012. *Alien Phenomenology: or What It's Like to Be a Thing*. Minnesota: University of Minnesota Press.

Bolter, J. D., and Diana Gromala. 2006. "Transparency and Reflectivity: Digital Art and the Aesthetics of Interface Design." In *Aesthetic Computing*, edited by Paul A Fishwick, 369–82. Cambridge, Massachusetts: MIT Press.

Bryant, L. 2014. *Onto-Cartography: an Ontology of Machines and Media*. Edinburgh University Press.

Cycling '74. 2015. "LOM – the Live Object Model." *Cycling '74*. Accessed October 11. https://docs.cycling74.com/max5/refpages/m4l-ref/m4l_live_object_model.html#Track.View.

Duignan, M. 2008. "Computer Mediated Music Production: a Study of Abstraction and Activity." PhD diss., Victoria University of Wellington.

Fell, M. 2013. "Collateral Damage: Mark Fell." *The Wire*. January. http://www.thewire.co.uk/in-writing/essays/collateral-damage-mark-fell.

Future Music. 2013. "Richard Devine Interview." *Music Radar*. http://www.musicradar.com/news/tech/richard-devine-interview-578399/.

Holmes, T. 1985. *Electronic and Experimental Music: Technology, Music and Culture*. 3rd ed. New York and London: Routledge.

Magnusson, T. 2010. "Designing Constraints: Composing and Performing with Digital Musical Systems." *Computer Music Journal* 34 (4): 62–73. doi:10.1162/COMJ_a_00026.

Maras, S.; and D. Sutton. 2000. "Medium Specificity Revisited." *Convergence* 6 (2): 98–113.

Meillassoux, Q. 2009. *After Finitude*. London, New York: Bloomsbury Publishing.

Mooney, J. 2010. "Frameworks and Affordances: Understanding the Tools of Music-Making." *Journal of Music, Technology and Education* 3 (2): 141–54.

Pold, S. B. 2011. "Interface Perception: the Cybernetic Mentality and Its Critics: Ubermorgan.com." In *Interface Criticism: Aesthetics Beyond the Buttons*, edited by Christian Ulrik Andersen and Søren Bro Pold, 91–113.

Aarhus University Press.

Power, B. 2007. “Power in the Box.” *Pro Sound News* 29 (12): 30.

Pras, A., C. Guastavino, and M. Lavoie. 2013. “The Impact of Technological Advances on Recording Studio Practices.” *Journal of the American Society for Information Science and Technology* 64 (3): 612–26. doi:10.1002/asi.22840.

Sterne, Jonathan. 2003. *The Audible Past: Cultural Origins of Sound Reproduction*. Durham: Duke University Press.

Sterne, J., ed. 2012. *The Sound Studies Reader*. New York and London: Routledge.

Stuhl, A. K. 2014. “Reactions to Analog Fetishism in Sound-Recording Cultures.” *The Velvet Light Trap* 74 (1): 42–53.

Waters, S. 2007. “Performance Ecosystems: Ecological Approaches to Musical Interaction.” In. Leicester, UK.

MUSIC SCREEN-READING: indicative results from two pilot studies

Lindsay Vickery
Edith Cowan University
WAAPA

Talisha Goh
Edith Cowan University
WAAPA

ABSTRACT

This paper discusses two pilot studies conducted by Lindsay Vickery and Talisha Goh, *Screening the Score: Exploring the Potentials and Limitations of Presenting Music Notation on the iPad* and *Sight-Reading Polyphonic Musical Textures: a Pilot Eye-Movement Study.* Vickery's experiment sought to investigate the activity of the eyes of musicians while performing a variety of notations and score presentation types from a screen. Goh's experiment explored sight-reading polyphonic keyboard music containing two, three, four and five voices, at a comfortable pace and with a click-track beat.

1. INTRODUCTION

Two pilot studies - *Screening the Score: Exploring the Potentials and Limitations of Presenting Music Notation on the iPad* (Lindsay Vickery) and *Sight-Reading Polyphonic Musical Textures: a Pilot Eye-Movement Study* (Talisha Goh) - were conducted examining the activity of the eyes of musicians while performing a variety of notation and score presentation types, and the effect of click-track speed on reading increasingly polyphonic music.

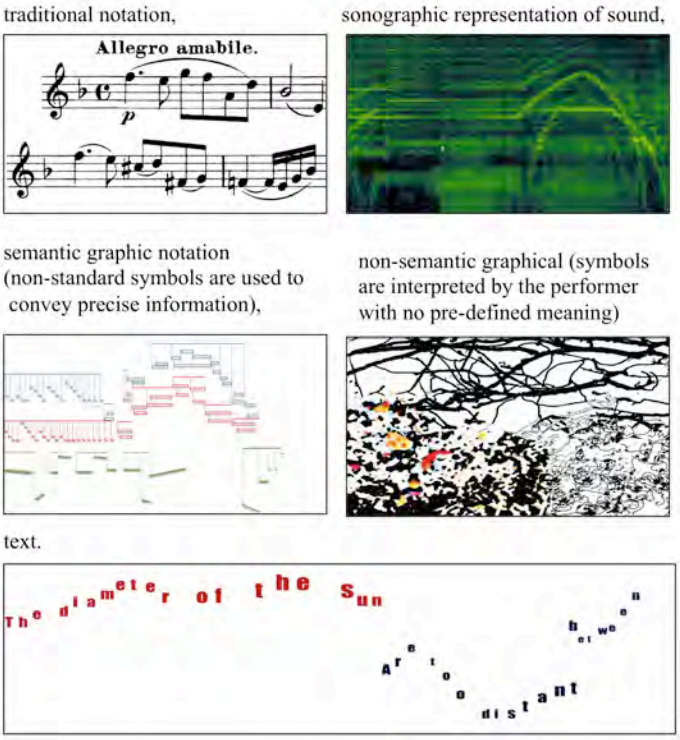


Figure 1. Traditional notation, sonographic representation of sound, semantic graphic notation, non-semantic graphical and text.

The goals of Vickery's study were to use eye-tracking technology to develop a methodology for investigating the effect of reading broad range of musical notation and presentation types from a computer or notepad screen and to identify potential anomalies for study more controlled examination. 11 subjects performing on a range of musical instruments (flute, clarinet, bass clarinet, violin, viola, piano, guitar, cimbalom and percussion) were tested, reading traditional notation, sonographic representation of sound, semantic graphic notation (in which non-standard symbols are used to convey precise information), non-semantic graphical (in which graphical symbols are interpreted by the performer with no pre-defined meaning) and text (Fig. 1.)

A range of screen-score presentation modes were also examined: traditional music score, scrolling score (moving from right to left across the screen), 2D scrolling score (moving both right to left and up and down across the screen), rhizomatic score (moving along predetermined 2D pathways) and scatter display (in which notation is presented in different parts of the screen) (Fig. 2.)

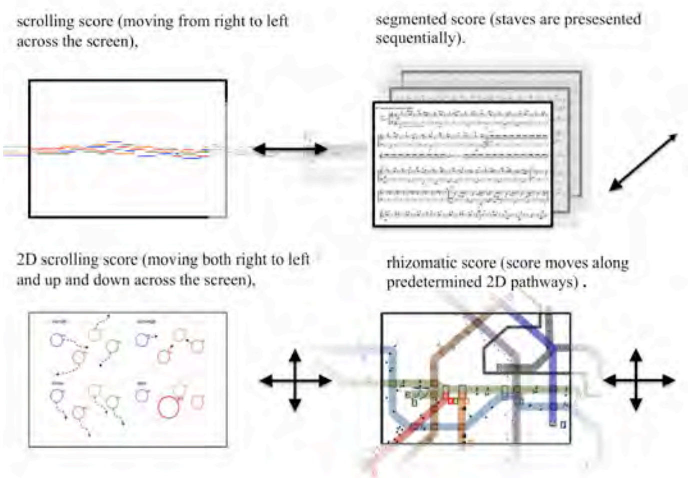


Figure 2. Traditional music score, scrolling score, 2D scrolling score, rhizomatic score and scatter display.

The goals of Goh's study were to examine the effect of increasing the polyphonic density of music notation and performing in conjunction with a click-track, upon performer's eye movements. In the experiment the eye movements of an expert musician were examined while sight-reading polyphonic keyboard music containing two, three, four and five voices, at a comfortable pace and with a click-track beat during three encounters with the music. The first (A) and fifth (E) examples are shown in Fig. 3 and Fig. 4 respectively. The number of fixations per second of performance, fixation duration, eye-hand span and position of fixations were assessed.

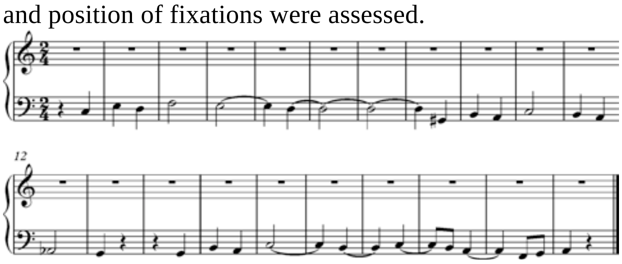


Figure 3. Excerpt A: Bass Voice

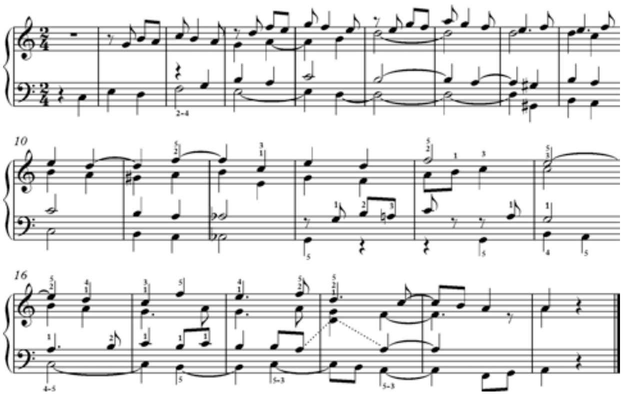



Figure 4. Excerpt E: Soprano, Alto, Tenor 1, Tenor 2 and Bass Voices

The technological setup for this study was supplied by Edith Cowan University's *Office of Research and Administration*, and consisted of a laptop computer connected to a large widescreen monitor (1920x1080 resolution; picture size 1900x900 pixels) and a web camera. The computer ran the program Tobii Studio v.3, which was designed to set up eye-tracking experiments and record and analyse the data acquired by a Tobii X2-30 eye-tracker was connected to the computer by USB (Tobii X2-30, 2014).

In addition to recording raw-data exportable in .xls format, the tracker visually represents fixation and saccade activity in real-time via video and allows the export of images of the cumulative plotting of fixation and saccade – "gaze-plot" - activity in relation to the score and "heat-maps" frequently viewed regions of the screen (Fig. 5).

Vickery's experiments were conducted in the Music Auditorium of the Western Australian Academy of Performing Arts (WAAPA). The musician's read from either image files of traditional notation or .AVI video files of "animated" scores presented on screen at a comfortable reading distance.

Goh's experiments were conducted in a designated room at Edith Cowan University. The computer screen with the eye-tracker was positioned in front of a digital keyboard with 88 weighted keys. It was also connected to a small damper foot pedal, again of comparable size to that of a piano.

The slides used for the experiment were generated using the Sibelius 7 manuscripts. These were converted into picture files that were compatible with Tobii Studio, and the music was displayed at a size comparable to a paper score. The slides were designed to display the excerpts in alphabetical order so that a voice was added on each successive excerpt.

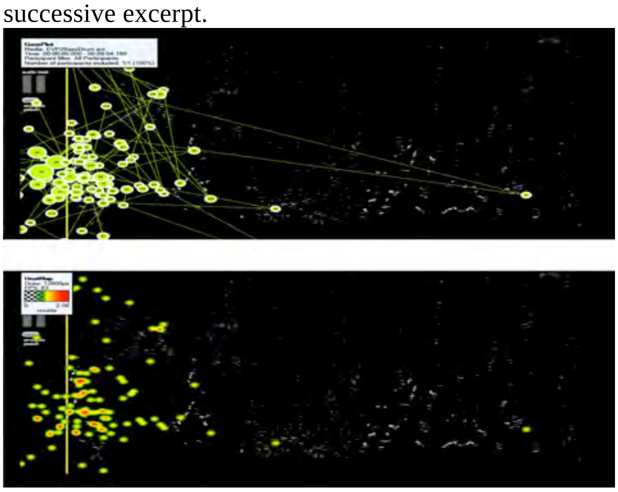


Figure 5. Gaze Plot (above) and Heat Map (below) of a Sonographic Score produced by the X2-30 eye-tracker.

2. CONTEXT OF RESEARCH

The exact nature of the processes involved in music-reading are still a matter of debate. One of the more persuasive models, proposed by Kinsler and Carpenter comprises a tripartite process "encoder, processor, executive" (1995: 1455). The pace at which visual symbols are encoded and processed is a key issue bearing the effectiveness of a screen score. This model proposes that the sight-reading process begins with encoding through preliminary transformation of the patterns into neural activity by retinal and central mechanisms, which is processed through the interpretation of musical symbols (traditional or otherwise), and that information about pitch and duration is finally transformed into appropriate patterns of commands to the executant muscles.

Eye-tracking studies focus at the surface of the encoding phase of this process by observing where, in what order and for what duration the eye is focusing, and moving. Performers visually acquire music notation with a combination of fixation upon graphical features and rapid repositioning of the eye (saccades). Sight-reading studies of traditional music are in agreement that fixation durations, extracting information from the score, fall within the range of 200-400 ms (Goolsby 1994a; 1994b, Truitt et al. 1997, Waters et al. 1997, Waters and Underwood 1998, and Gilman and Underwood 2003). The durations of saccades between fixations while sight-reading fall within in the tens of milliseconds (Gilman and Underwood 203:221).

Each fixation takes in a region termed the gaze frame or perceptual span which has been demonstrated to extend for "approximately 1 measure right of the fixation point" (Truitt et al. 1997) (equivalent to 3-5cm). Contrary to popular belief, the extent to which performers read ahead of their execution, termed the Eye-hand Span is relatively small, being between approximately 2 and 4 beats (Rayner & Pollatsek, 1997). Even in skilled readers visual processing of notation is not very far ahead of the hands and the actual position in the score (Gunter et al 2003:742).

The extremely short durations of some eye-movements (in the range of 25-30 milliseconds (Rayner, 1978)) meant that early systems using film (Weaver 1943) were unable to capture subtle eye movements (Goolsby, 1994a). Since the mid 1970's, notably the release and availability of the home computer, technology has allowed the revival of eye-tracking research. Rayner (1998: 372) identified this as the 'third era' of research that enabled the collection of more accurate data and larger amounts of data. Goolsby (1994a)'s study is said to have established modern musical eye-tracking practice because it was the first to utilise a dedicated eye-tracking device which measured the horizontal and vertical eye position every millisecond (Goolsby, 1994a; Madell & Hébert, 2008). Contemporary systems including the one used in these studies generally use the reflection of infrared light off the cornea to determine the angle of the eyes.

The applicability of the significant literature exploring the mechanism of sight-reading of traditional musical notation is limited in a number of ways. The tasks in many sight-reading studies involve quite simple musical examples (especially in comparison to the works of many composers) and because of the ease of collecting accurate data from MIDI keyboards, sight-reading studies have typically focused upon keyboard players, and have therefore not taken into account variation in the performance of instrumentalists who must place musical notation at a significant distance: a keyboard player might typically read from a score at a distance of 50 cm, whereas a percussionist may need to place a score 150 cm away in order to allow for a large instrument or set up.

Very few studies specifically address the issue of reading music from screens. In 1997 Picking compared a number of presentation styles for musical notation including traditional paper-based medium, as well as its screen-based counterpart however the study used bit-mapped notation presented in hypercard stacks, now many technological generations of out of date, and was undertaken in an era when smart phone and tablet literacy was not ubiquitous.

Importantly there are currently no studies of reading of nonstandard musical notation. This is perhaps due to its degree of specialisation involved, as well as the idiosyncratic nature and variety of graphic notations. Similarly the effect of performer restraint via a click-track had also not been studied.

3. INDICTATIVE RESULTS

Both experiments were designed as pilot studies to identify issues for further investigation. The reasons that neither study was able to attain fully valid results are discussed in section 4. For this reason only the raw data and its potential implications are discussed.

3.1 Vickery: Screening the Score

The graphs in figures 6-11 show a multicoloured band representing the average result for each participant for each case and then a black bar representing the average result for that case.

Average fixation durations were compared between *notation* types and found to increase in the following order: traditional notation 647ms, sonographic representation 681ms, non-semantic graphical notation 803ms, text 965ms and semantic graphical notation 2604ms (Fig. 6.).

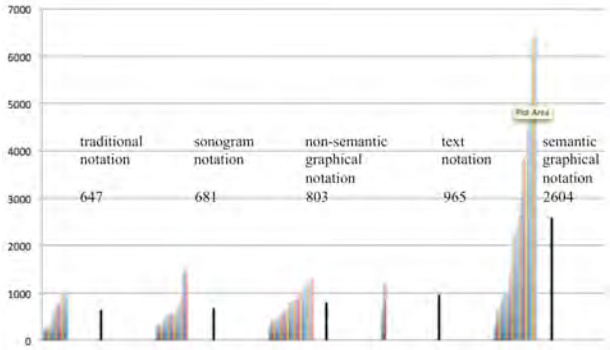


Figure 6. Average fixation length (ms) versus notation type.

These results may reflect that notation that requires the performer to continuously follow the evolution of graphical shapes require greater fixation durations and that graphical shapes with specific semantic meanings require greater fixation durations than graphical shapes that allow greater interpretation by the performer. Average fixation durations were compared between score *presentation* types and found to increase in the following manner: traditional score 566ms, 2D scrolling 819ms, rhizomatic 823ms, scatter display 893ms and scrolling score 1560ms (Fig. 7.).

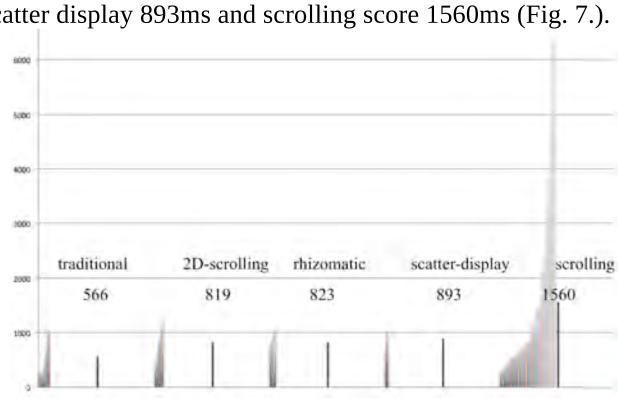


Figure 7. Average fixation length (ms) versus presentation type.

However, the range between the lowest and highest of fixation lengths in each category (staff 785ms, 2D scrolling 804ms, rhizomatic 279ms, scatter display 400ms, scrolling 6147ms) suggest that notation type may play a more important role than presentation type in determining fixation length.

Saccades, the movements of the eye between fixations, were also examined. Average saccade durations were compared between *notation* types and found to increase in the following manner: semantic graphical notation 86.5ms, text 99ms, staff notation 108.25ms, sonographic representation 120.8ms and non-semantic graphical notation 122.1ms (Fig. 8.). When average saccade durations were compared between score *presentation* types it was found that average fixation durations increased in the following manner: rhizomatic 92ms, scatter display 96ms, 2D scrolling 97ms, traditional score 108ms, and scrolling score 113ms (Fig. 9.). Again the data suggests that notation type has a more important effect upon saccade duration than presentation type.

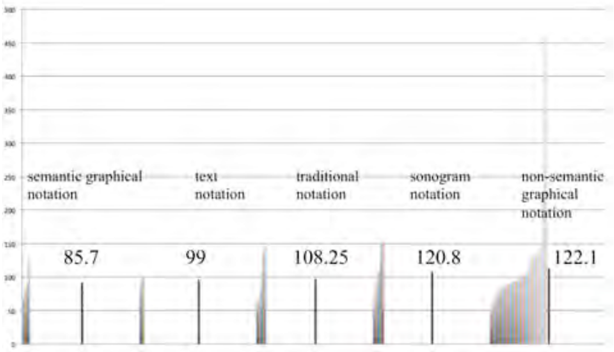


Figure 8. Average saccade length (ms) versus notation type.

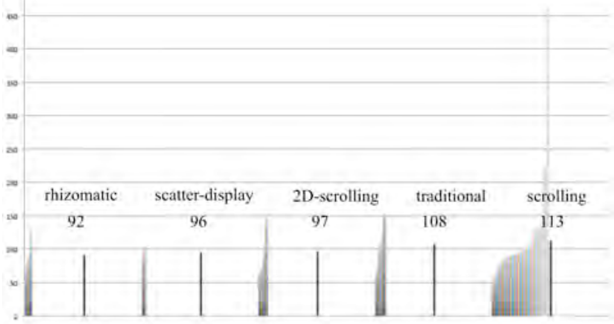


Figure 9. Average saccade length (ms) versus presentation type.

The ratio (%) of the total number of fixations versus the total number of saccades was also examined. Amongst *notation* types the ratio increased in the following manner: traditional 72.2%, non-semantic 72.7%, spectrogram 78.9%, semantic 83.9%, text 85.7% (Fig. 10.). Amongst *presentation* types the ratios found were: traditional 66%, 2D scrolling non-semantic 76%, rhizomatic 78%, scrolling 79% and scatter display 82%, (Fig. 11.). The results showed more fixation activity in non-traditional scores and the highest activity in scatter displays in which the notation was changing in multiple sectors of the screen.

Significant variation was found in the width of participants scan patterns: the broadest gaze point width was for a Sonogram-style notation work, with a standard deviation of 272 px (X axis) and 224 px (Y axis); the narrowest was for a the semantic graphical notation work, with a standard deviation of 100 px (X axis) and 56 px (Y axis) (Figure 13.).

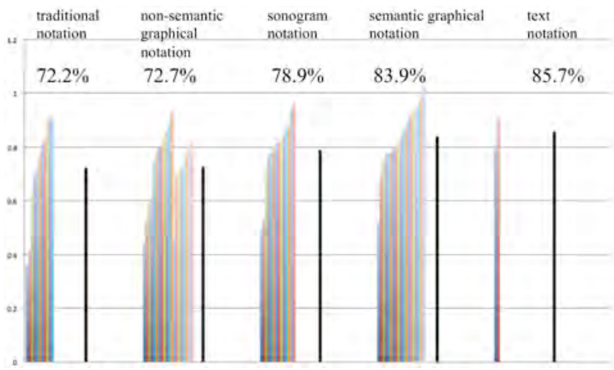


Figure 10. Fixation number to saccade number ratio (%) versus notation type.

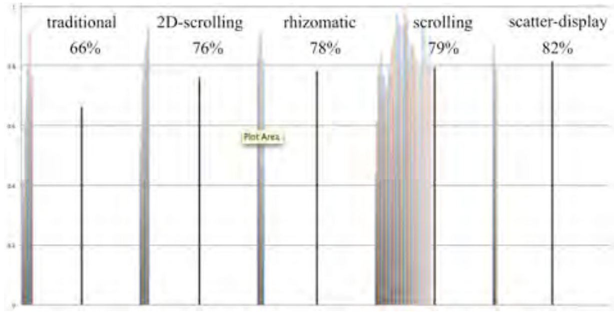


Figure 11. Fixation number to saccade number ratio versus presentation type.

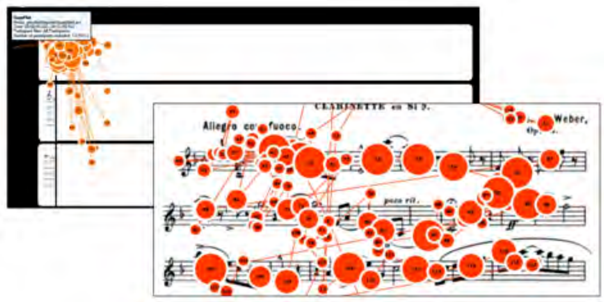


Figure 12. Comparison of Scrolling score gaze-plot in which the eye is restricted to a small area (above) and traditional notation gaze plot (below) in which the eye scans across staves.

3.2 Goh: Sight-Reading Polyphonic Musical Textures

The most significant effect was seen in reading with a click-track, which decreased the number of fixations per second and eye-hand span, and increased the average fixations duration. This was attributed to a higher cognitive load involved in performing with a click-track. Musical texture also interacted with eye movements. Polyphonic excerpts containing more voices correlated with a higher number of fixations per second and lower eye-hand span and fixation durations. Unusual musical features accounted for abnormally large numbers of fixations and time spent looking at areas of the score.

The principle findings of the results were that eye movement patterns were affected by the number of voices in the excerpt and the click-track condition, but not the number of encounters. Furthermore, the fixation positions were influenced by features of the music and notation.

The click-track condition performances were markedly slower and more difficult for the participants than the comfortable condition. This was reflected in the average fixation duration, number of fixations per second and eye-hand spans of the excerpts. The average number of fixations and fixation duration were notably higher when participants performed with a click-track. However, when accounting for the time it took to perform the excerpts the number of fixations per second of performance was actually lower in the click-track condition. In addition, the eye-hand span was markedly smaller when playing with the click-track (1.36 beats compared to 1.86 when played comfortably.)

Excerpts containing more voices had slightly more average fixations, more progressive fixations and shorter average fixation durations than excerpts with less voices.

The number of fixations per second was positively correlated with the number of voices in each score (Fig. 13).

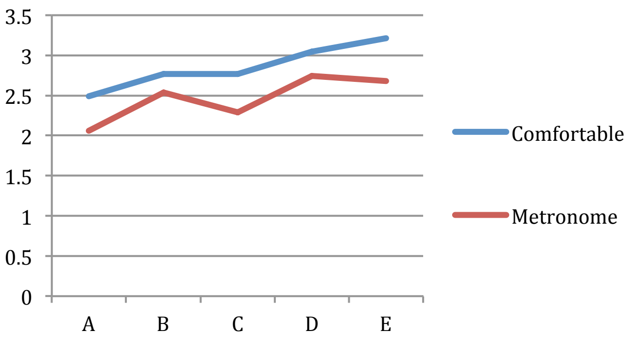


Figure 13. Number of Fixations per Second.

Fixation duration was slightly negatively correlated with the number of voices of each excerpt (Fig. 14.) The eye-hand span was negatively correlated to the number of voices in each excerpt in the comfortable condition, but not in the click-track condition.

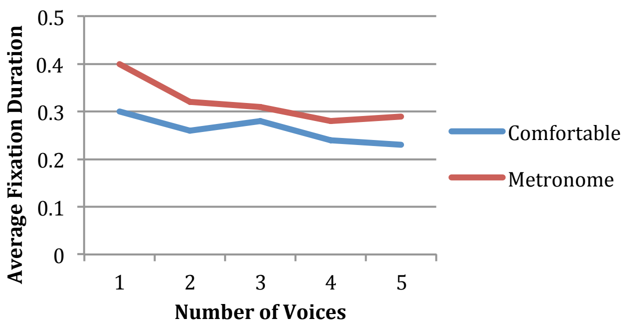


Figure 14. Average Fixation Duration Across Excerpts.

In general, fixations were made within a stave but not on note-heads themselves. Fixations tended to occur between staves in the two-handed excerpts, and underneath the bass stave in the one-handed excerpt. Subjects tended to read in a horizontal fashion, a trend reported in Weaver (1943) for contrapuntal music. Overall, the fixation positions suggest that musicians rely heavily upon perceptual span and peripheral vision. This could be an element that is important in chunking behaviour.

In Excerpt A, which only involved the left hand, fixations tended to occur slightly below the note heads themselves (Fig. 15.). Excerpt E shows a pattern of focus between the treble and bass staves rather than on any one line itself (Fig. 16.). Although this could be due to a calibration error, a more likely explanation is that peripheral vision was important in reading these excerpts. Peripheral vision and the perceptual span have an important role in chunking groups of notes, rather than focusing on one note at a time. Other studies, such as Gilman and Underwood (2003) and Truitt et al. (1997), have also commented on the proportion of blank space that was fixated upon. Although these studies did not specify where the 'blank space' was, the current study has also found a focus on the empty space between staves and underneath the staff in the one-handed excerpt.

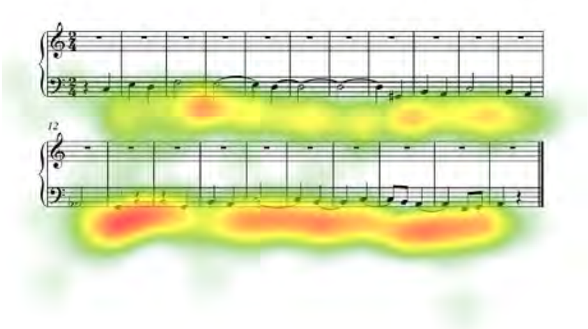


Figure 15. Heatmap: Excerpt A, Number of Fixations.

Another striking feature of the fixation positions was the patterns created by sequential fixations. Although the fixations were rarely on the notes themselves, especially those in the treble clef, the eyes tended to move according to the main features of each bar of music.

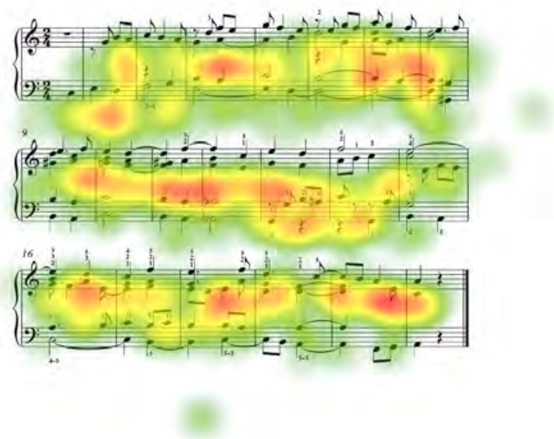


Figure 16. Heatmap: Excerpt E, Number of Fixations.

4. LIMITATIONS OF THE EXPERIMENTS

Vickery's study was deliberately broad in order to identify issues for further investigation. While it was arguably successful in this regard there were too great a number of independent variables in the study to make conclusive reliable findings. Sight-reading studies suggest that the baseline for the duration of fixations on musical symbols is between 200-400 ms per fixation (Goolsby 1994a; 1994b, Truitt et al. 1997, Waters et al.

1997, Waters and Underwood 1998, and Gilman and Underwood 2003). The average fixation length of the participants in this study for traditional notation was 556ms with a range between 272 and 1057ms. Average fixation durations amongst participants in this study ranged from 450ms (player 1) to 2089ms (player 8). It is not clear whether this accurately reflects the reading style of the participants or the difficulties of standardising data acquisition across such a wide range of instruments. For example Infrared scanning may have been affected by the distance between the performer and the apparatus and/or the reflectivity of some instruments.

Goolsby (1994) has suggested that notational complexity is a strong determinant in predicting eye-movement in music) and this effect was not compensated for. Future work should include the development of means of evaluating notation complexity, in order to allow accurate comparison between notation types. Although Vickery has speculated on the role of scroll speed in scrolling scores (Vickery 2014), this issue was also not compensated for.

Susan George claims that “in the most general sense the score is comprised of units. Sometimes these units are primitive elements themselves, and sometimes they are composite so that the primitive elements must be extracted from the units themselves” (2004: 157). In website analysis these are referred to as "semantically meaningful units", and used to measure the interaction of a reader with the screened page. Such an approach (Fig. 17), might usefully be adopted in the study of performers’ interaction with screenscores, as a means of measuring the number of semantically meaningful units that a performer is able to capture in a single fixation and the rate at which they can be captured.

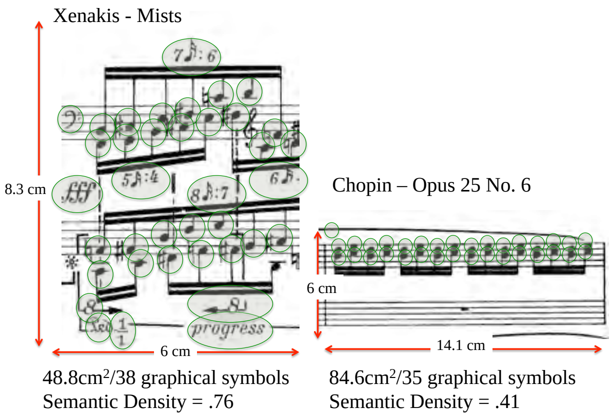


Figure 17. Calculating Semantic Density by measuring graphical symbols per cm^2.

In Goh’s study the small number of participants restricts the validity of the results. The most unusual finding – that the subjects did not generally improve playing performance across encounters - is unusual, as it is contrary to the gradual improvement in perceptual span across performances reported by Burman and Booth (2009) and the improvement in musical performance reported by Goolsby (1994a). A possible explanation is that the excerpts were too simple for the participants for any improvement to be recorded, or that the three performances analysed were insufficient to record improvements in performance. Future studies should focus on a more longitudinal approach and a greater variety of musical textures to observe the effects of rehearsal on music reading performance.

5. DISCUSSION

Given the limitations upon the validity of the findings suggested above, the results suggest that:

1. a score’s Notation type has a greater influence on eye-movement in music reading than Presentation type.

2. participants fixation length was shortest when reading traditional notation, increased through more interpretive forms of non-semantic graphical notation to a peak when reading semantic graphical notation. This may be the result of the fact traditional notation consists of a symbolic language that is read in discrete chunks, whereas graphical notation implies more continuous monitoring of notational changes and that semantic graphical notation, the most precise form of graphical notation demands the most continuous attention and therefore has the longest fixation lengths.

3. Semantic graphical notation had the shortest average saccade time and non-semantic graphical notation the longest. Similarly, the ratio of “unclassified gaze event types” was shortest for semantic and longest of non-semantic forms of notation and the scan pattern width was also smallest for semantic notation. These findings support the notion that semantic notation requires the eye to remain relatively fixed in a confined area, whereas non-semantic graphical notation encourages a more open, interrogative approach to reading.

4. performance with a click-track, decreased the number of fixations per second and the eye-hand span. This was thought to be an effect of the high cognitive load involved in playing with a click-track. There are many other factors that can influence musical sight-reading like psychomotor speed and general cognitive skills (Kopiez & In Lee, 2008), and so there may be other explanations for this phenomenon.

5. A correlation was observed between musical texture and eye movements as polyphonic excerpts with higher numbers of voices tended to produce a smaller eye-hand span and more fixations per second of performance, and these fixations lasted a shorter duration. As no other studies have focussed on polyphonic textures, this investigation is the first to report this effect.

6. CONCLUSION

The (slightly miraculous) process by which performers encode, process, execute notated musical structures in an extremely precise temporal frame may never be fully understood. These studies aimed to shed light on some aspects of critical importance to composers seeking to coordinate performers through computer control. They examined the effects upon eye-movements of a range of notation and presentation types in score-reading and the sight-reading of polyphonic musical textures with a click-track beat through several encounters. A number of interesting results were recorded and require further isolated study in order to develop reliable and valid data. Additional studies will need to be conducted to see if these findings apply to a wider population and to make valid generalisations from the results.

Although the findings of Goh's study agree with the current sight-reading literature, future studies should incorporate multiple tempo conditions to fully examine the effect that click-track playing has on the cognitive load.

As eye-tracking technology becomes more advanced and accessible, music sight-reading studies will reach beyond the basic phenomena in this fascinating field. This study has explored the effects forms of notation and score presentation, polyphonic textures, multiple encounters, cognitive load and musical features on sight-reading. Future studies should take all aspects of the sight-reading context into consideration in order to explain eye movement data. It is hoped that such studies will provide solid and beneficial data for the increasing number of composers working with coordination of performers through computer via notation on screen and/or click-track.

7. REFERENCES

Burman, D. D., & Booth, J. R. 2009. Music rehearsal increases the perceptual span for notation.

George, S. E. 2004. Visual Perception of Music Notation: On-line and Off-line Recognition. IRM Press: Hersey, PA.

Gilman, E., & Underwood, G. 2003. Restricting the field of view to investigate the perceptual spans of pianists. Visual Cognition, 10(2), 201-232.

Goolsby, T. W. 1994a. Eye movement in music reading: Effects of reading ability, notational complexity, and encounters. Music Perception, 77-96.

Goolsby, T. W. 1994b. Profiles of processing: Eye movements during sightreading. Music Perception, 97-123.

Gunter, T. C., B.-H. Schmidt, et al. 2003. Let's face the music: A behavioral and electrophysiological exploration of score reading. Psychophysiology 40: 742–751.

Kinsler, V., & Carpenter, R. 1995. Saccadic eye movements while reading music. Vision Research, 35(10), 1447-1458.

Kopiez, R., & In Lee, J. 2008. Towards a general model of skills involved in sight reading music. Music Education Research, 10(1), 41-62.

Madell, J., & Hébert, S. 2008. Eye Movements and Music Reading: Where Do We Look Next? Music Perception: An Interdisciplinary Journal, 26(2), 157-170. doi: 10.1525/mp.2008.26.2.157

Picking, R. 1997. Reading Music From Screens Vs Paper, Behaviour & Information Technology, 162:72-78

Rayner, K. 1978. Eye movements in reading and information processing. Psychological Bulletin, 85(3), 618.

Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124(3), 372-422. doi: 10.1037/0033-2909.124.3.372

Rayner, K. 2009. Eye movements and attention in reading, scene perception, and visual search. The Quarterly Journal of Experimental Psychology, 62(8), 1457-1506. doi: 10.1080/17470210902816461

Rayner, K., & Pollatsek, A. 1997. Eye Movements, the Eye-Hand Span, and the Perceptual Span During Sight-Reading of Music. Current Directions in Psychological Science, 6(2), 49-53. doi: 10.1111/1467-8721.ep11512647

Tobii X2-30 Eye Tracker User's Manual. 2014 Retrieved 1 November, 2014, from http://www.tobii.com/Global/Analysis/Downloads/User_Manuals_and_Guides/Tobii_X2-30_EyeTrackerUserManual_WEB.pdf

Truitt, F. E., Clifton, C., Pollatsek, A., & Rayner, K. (1997). The Perceptual Span and the Eye-Hand Span in Sight Reading Music. Visual Cognition, 4(2), 143-161. doi: 10.1080/713756756

Vickery, L. 2014. The Limitations of Representing Sound and Notation on Screen. Organised Sound 19(3). Cambridge University Press.

Waters, A. J., & Underwood, G. 1998. Eye Movements in a Simple Music Reading Task: A Study of Expert and Novice Musicians. Psychology of Music, 26(1), 46-60. doi: 10.1177/0305735698261005

Waters, A. J., Underwood, G., & Findlay, J. M. 1997. Studying expertise in music reading: Use of a pattern-matching paradigm. Perception & psychophysics, 59(4), 477-488.

Weaver, H. E. 1943. Studies of ocular behavior in music reading. Psychological Monographs, 55(1), i-50. doi: 10.1037/h0093537

THROUGH THE EYE OF THE NEEDLE: COMPOSITIONAL APPLICATIONS FOR VISUAL/SONIC INTERPLAY

Lindsay Vickery
Edith Cowan University
WAAPA

ABSTRACT

This paper explores the relationships between sound and its visualisation, focusing upon the issues of representation and interpretation of music through both performative and machine processes. The discussion proceeds in the context of recent works by the author exploring the representation of sound and musical notation and their relationship to and with performance: *murmurs trapped beneath the bark* [2014], *in nomine tenebris* [2014], *detritus* [2015], *trash vortex* [2015], *acid fury* [2015], *here, apparently, there was time for everything* [2015], *between-states* [2015], *the miracle of the rose* [2015], *residual drift* [2015], *...with the fishes...* [2015] and *inhabited matter* [2015]. Issues examined include: resonification of spectrograms, visualisation of spectral analysis data, control of spatialisation and audio processing using spectral analysis data, and reading issues related to scrolling screen score notation.

1. INTRODUCTION

In 2013, through exploration into the development of notation that would be capable of capturing the nuances of complex sounds, I became interested in the possibility of using the spectrogram of a sound, more or less annotated, as the source for musical notation. The immediate question was what resolution would a spectrogram need to be in order to capture all the nuances that the ear can detect?

The eye and the ear sense vision and sound in entirely different ways. In the simplest sense, putting aside the processing that occurs at a later point to create the auditory scene, sound is mapped in a broadly linear fashion with the cochlea capturing frequencies from high to low with a resolution of roughly 20ms. The eye in contrast combines data from a range of different sensors – colour through three cone cells and luminosity through rod cells. The result is that vision is not mapped in a linear fashion: if it were, the light spectrum would appear as a continuum from purple - the brightest (highest) colour to red the darkest (lowest colour). The arrangement of rods and cones gives rise to anomalies such as the non-sequential perceptual "brightness" of colours such as yellow, cyan and magenta in the colour spectrum. Additionally, in order to focus on a score with a "gaze frame" of roughly 4cm^2 the eye must fixate the fovea – the only part of the eye densely enough packed with cones to show the required detail – for periods of time in the order of 200 - 400ms. Sound is akin to a rapidly changing column and vision to a slowly changing field.

Using this logic, in order to attain the detail that is sonically perceptible, a visual score would hypothetically need to be changing at a rate of approximately 40cm/s. (My previous research has suggested reading becomes uncomfortable beyond a rate of approximately 3cm/s (Vickery 2014b).

Musicians are capable of performing nuances at extremely minute durations, but between the sound and the performance lies the notation. I have come to refer to this restriction as "the eye of the needle".

Peter Ablinger explores this issue in his "Phonorealist" works such as the *Quadraten* series, in which spectral analysis data from recordings is "reconstituted in various media: instrumental ensembles, white noise, or computer-controlled player piano" (Barrett 2007). A key issue at the heart of *Quadraten* is representation or analogy made between "real" sounds and their reconstituted counterparts. The "eye of the needle" problem obliges a loss of resolution and this loss can become interesting in itself.

> The reproduction of "phonographs" by instruments can be compared to photo-realist painting, or - what describes the technical aspect of the "Quadraturen" more precisely -with techniques in the graphic arts that use grids to transform photos into prints. Using a smaller grain, e.g. 16 units per second, the original source approaches the border of recognition within the reproduction. (Ablinger 2011)

This paper examines some developments in this work exploring the relationships between sound and its visualisation, focusing upon the issues of representation and interpretation of music through both performative and machine processes. Some initial projects exploring this issue were discussed in an earlier paper (Vickery 2014). The discussion proceeds in the context of recent works by the author exploring the representation of sound and musical notation and their relationship to and with performance including *in nomine tenebris* [2014], *detritus* [2015], *trash vortex* [2015], *acid fury* [2015], *between-states* [2015], *the miracle of the rose* [2015], *residual drift* [2015] and *...with the fishes...* [2015], that have resulted in new approaches to notation and new possibilities for composition and sound processing.

2. SONIFICATION/VISUALISATION/ (MANIPULATION)/RESONIFICATION

In 2013 I began a project to construct means to interchange data between visual and sonic media: to create a continuum in which sound could be visualized and then resonified. A patch, *Sinereader,* was created in *MaxMSP* to resonify greyscale spectrogram images (Fig. 1). In the patch each vertical pixel of a greyscale version of the spectrogram of a sound is mapped to one of 613 independent sinewave generators. In the patch a .png file of the spectrogram is loaded into a jit.qt.movie, it is then played through jit.matrix and jit.submatrix that send an image of one pixel width to the jit.pwindow. Data from the submatrix is split into a list of 613 values in jit.spill and these values are represented in a multislider. The vertical pixels are scaled logarithmically according to the vertical resolution of the spectrogram and each mapped to an individual cycle~ object. The greyscale value of each pixel is scales and mapped to the amplitude of each `cycle~` object. In addition to being a transcription tool, the patch can also be controlled externally as an "instrument" using *MaxMSP Mira* app for iPad.

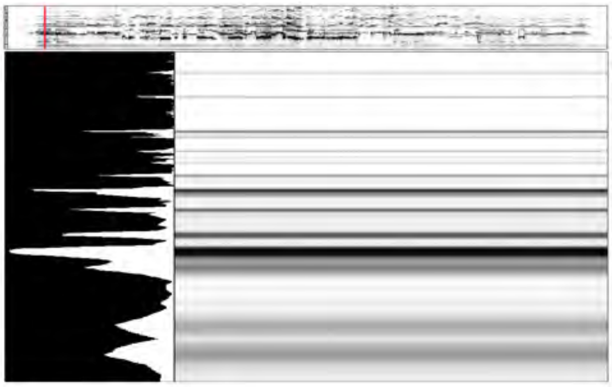


Figure 1. The *Sinereader* patch developed in *MaxMSP* to resonify spectrograms. The complete spectrogram with a "scrollbar" indicating progress through the image is displayed at the top of the image, the greyscale value of each vertical pixel in a one pixel segment is displayed on the bottom right and the resulting amplitude is displayed on the bottom left.

This particular form of resonification is not the only possible route and even in the case of the Sineplayer itself there are other potential end points that have not yet been explored: for example sonification through granular or wave table synthesis. However, even within this fairly straightforward resonification process there is much potential for exploration.

In *in nomine tenebris* [2014] a new step was added to this process: a spectrogram of Giacinto Scelsi's organ work *In Nomine Lucis* [1974] was processed in the visual domain using the software *Photoshop* and *Illustrator* to stretch, warp and distort the spectrogram image (Fig. 2). The image was then resonified and used as the basis for a score for five acoustic instruments. Scelsi's original quite static work provides a spectrally and somewhat morphologically defined sonic substance from which the new work is fashioned, new sonic morphologies are wrought on the material, formal proportions are manipulated and a new work is created sharing a kind of spectral DNA with the original.

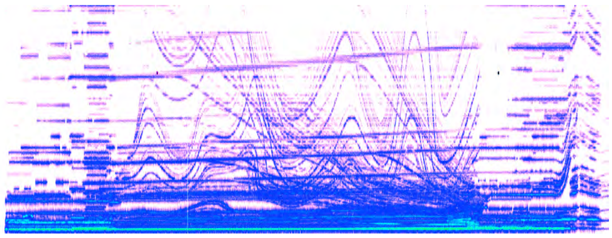


Figure 2. An excerpt of the manipulated spectrogram of Scelsi's *In Nomine Lucis* used to create *in nomine tenebris*.

acid fury [2015] explores what O'Callaghan terms "Category 2 Musical Mimesis" through the transduction of a found sound recording (of chaotic noise and a distant power tool) into a sound object (and score) of "some acoustic similarity to the source sound, but distant enough that it requires other extra-musical contexts to identify" (O'Callaghan 2012). Again the spectrogram is re-synthesised using the *Sinereader* patch to produce a sound object morphologically almost identical to the original recording, and yet distinct in an acousmatic sense.

Between these two approaches, *residual drift* [2015] (a homage to the 50th anniversary of Varèse' death), takes a spectrogram of a recording of *Ionisation* that was processed in *Illustrator* to generate a score and accompanying pre-recorded electronics. Spectrograms are probably least effective at visualising percussive works, and this deficiency was exploited by choosing a resolution for the spectrogram that would render features such as snare drum rolls as continuous pitches. Because the spectrogram was restricted to a limited frequency band (the range of the bass flute C2-C4) its resynthesised audio was reprocessed by analysing the recording using the `sigmund~` object and using the detected pitches to generate "false harmonics" by ring modulating the audio multiple times to add higher but related upper frequencies.

The final version of the electronics part incorporates three versions of the sonification: processed by *Photoshop*, processed by *Illustrator* with added false harmonics and a version featuring a spectrogram of just the upper harmonics of the recording (which were excluded in the image used to make the score). The "high harmonics" version prominently features the imprint of *Ionisation*'s characteristic sirens.

3. NOTATION FROM SPECTROGRAMS

Two important consequences of the emergence of Electroacoustic music upon instrumental music were the development by Lachenmann of "Musique Concrete Instrumentale"

> (t)he term refers to that of Pierre Schaeffer's musique concrete. Instead of instrumentally using mechanical everyday actions as musical elements I go about it by understanding the instrumental sound as information

> about its production, therefore rather the other way round - by illuminating instrumental sounds as mechanical processes. (Lachenmann 1996: 211-212)

and by Grisey of “instrumental synthesis”

> Grisey calls ‘instrumental synthesis’, where a recorded sound is orchestrated based on an analysis of its frequency content over time (either by the visual aid of a spectrogram or, as is often the case in contemporary efforts, increasingly sophisticated and diverse software analysis tools). (O’Callaghan 2015)

These two expansions of the language of instrumental music opened new fields of possibilities, but also new problems for the representation of complex sounds for live performers. Lachenmann’s solution, for example in *Pression* [1966], was often to turn to tablature – that is to define the actions of the performer rather than the sound to be produced. Grisey often retained aspects of traditional notation but adopted the proportional duration approach of Earle Brown and Berio (for example in *Partiels* [1975]). Both options have trade-offs in precision and neither solve the underlying problem, that there are limitations upon how much notational detail can be delivered in real-time to a musician.

Aaron Cassidy’s increasing unease with the complexity of his gesturally polyphonic tablature style scores, (such as *What then renders these forces visible is a strange smile…* [2007-8]) has led to his proposing a “unified multi-parametric notation system” in which “multi-planar, multi-dimensional movements are reduced to a two-dimensional notational image (principally through the use of color)” (Cassidy 2013).

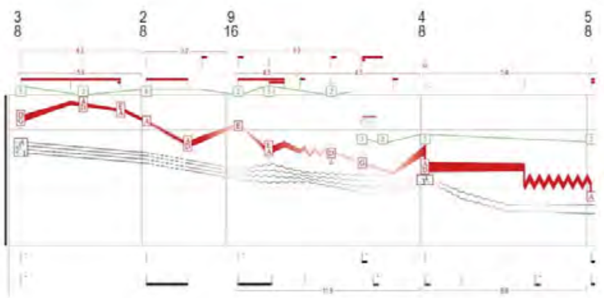


Figure 3. *Aaron Cassidy’s* “unified multi-parametric notation system” *in Second String Quartet*, Violin I, mm.1-4.

Another solution might be to draw from the discipline of cartography and create digital scores that can be viewed in multiple modes (in the way a map can be viewed in satellite or terrain mode) allowing the composer to define the sound in one mode and the action in another. Modern cartography has also developed means to “zoom-in” on details while maintaining a fairly constant graphical density, for example using software to display no more than 10 “semantically meaningful units” per cm^2 (Bertin 1967). Such a system would allow a performer to “zoom in” on detail of a complex notational passage, commit it to “muscle memory” and then “zoom out” to a level of notational signification that is actually readable at speed.

My own recent work has tended to focus the representation sound to be produced by the musicians rather than the means. An early approach to this problem was *unhörbares wird hörbar* [2013] in which a score highlighting important sonic features was created from a “spectral trace” in which the notation was drawn directly onto the spectrogram. In *in nomine tenebris* [2014], parts for five instruments were created by directly colouring and annotating features of the spectrogram itself (Fig. 4.). This allowed performers to acoustically recreate minute and continuous changes in pitch, dynamic and timbre by following the height, thickness, shape and shade of the original spectrogram. The players performed the work from scrolling scores synchronised in the *Decibel Scoreplayer* (Wyatt et al. 2012).

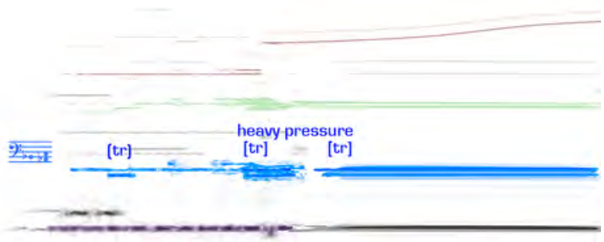


Figure 4. The opening of the cello part *for in nomine tenebris* [2014].

In *acid fury* colour-coded parts for the eight instruments were made from direct transcriptions of a spectrogram of the recordings (Fig. 5.). Here the correspondence between the recording and the instrumental lines is extremely difficult to perceive as they are extremely rapid and simultaneous – the “noisy” sections (it actually sounds reminiscent of Ornette Coleman’s *Science Fiction* [1971]) and the effectiveness of the work is almost entirely based on the precise coordination of the transitions between chaos and relative stillness facilitated by the networked scrolling score.

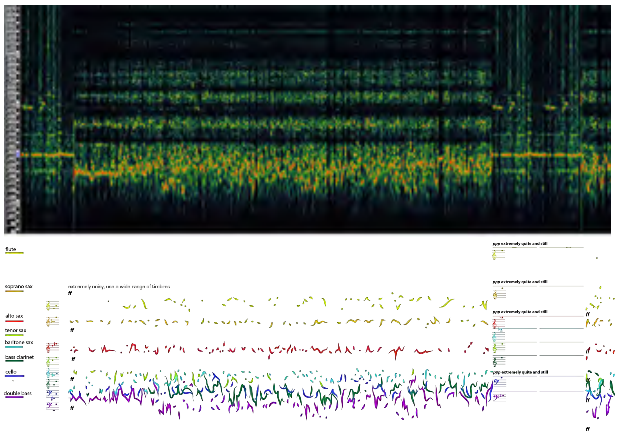


Figure 5. Spectrogram of found-sound recording (above) and annotated graphical transcription (below) of *acid fury* [2015] (excerpt).

Sonic Visualiser software offers the possibility of annotating a spectrogram directly in a MIDI layer. In the work *here, apparently, there was time for everything*

[2015] the initial aim was to create a notated score created from a MIDI transcription of a spectrogram of "no-input bass clarinet"[1] feedback that was processed by complex variable-speed delay. The process of “pruning” the MIDI transcription and “orchestrating” specific pitches to instruments (soprano saxophone, electric guitar and harp) was begun (Fig. 6), however the problem of imposing a temporal grid to the unmetered music and the desire for the musicians to perform in synchronisation with the source recording led to a return to the “spectral trace” approach to transcription.

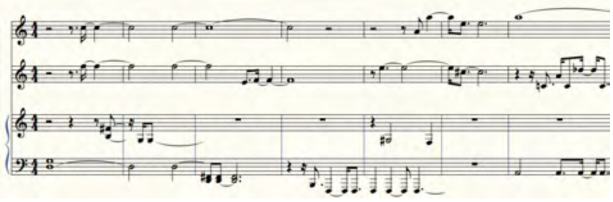



Figure 6. “Orchestrated” MIDI transcription of *here, apparently, there was time for everything* [2015] (excerpt).

This work, featuring textures far closer to those accommodated by traditional notation than *unhörbares wird hörbar* introduced a new problem: horizontally, the spectral trace is temporally proportional and performing with precision can be achieved through a scrolling score, however, it is also vertically proportional – each pitch occupies a distinct vertical spatial position. Because the source audio had no glissandi and was less timbrally divergent than my previous spectral trace works, a vertically proportional stave was the most appropriate to represent the pitches. The issue of representing the chromatic scale on a staff has occupied many minds for some time (Morris u.d), and resulted in systems using varied lines, spacings and noteheads. The system I developed attempt to more-or-less retain the topographical layout of the traditional stave, while adding coloured lines to accommodate non-natural notes (Fig. 7).

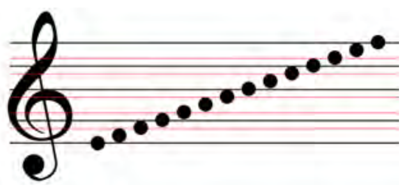


Figure 7. Proposed vertically proportional stave showing a chromatic scale from E3 to F4.

Fig. 8 is the opening of the scrolling transcription of *here, apparently…* in addition to the proportional stave, note stems/beams are used as an indicator of phrase grouping and circular noteheads are adopted. Stems are always located to the left of noteheads as it aids in reading synchronously in a scrolling score. Although I personally prefer the proportional stave version, for the first performance – due to the small number of rehearsals - a scrolling non-proportional stave version was also created.

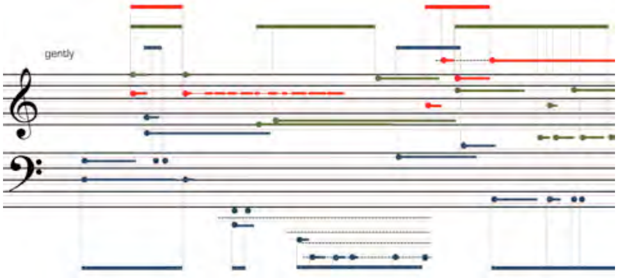
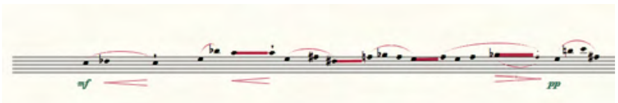


Figure 8. Vertically proportional stave transcription of *here, apparently, there was time for everything* [2015] (excerpt).

Despite a certain elegant schematic quality to the score, I still feel the level of abstraction from the spectrogram’s descriptive shapes and timbres, removes too much musical information for the performers.

In *between-states* [2015] the initial concept was also to construct a traditionally notated score for bass flute and bass clarinet from a spectrogram of a mosaic of transformations of acoustic bass flute and bass clarinet recordings. The transformations were then edited leaving the most interesting outcomes still in their original temporal position - and therefore all relating back to the source acoustic recording. The process of creating a traditionally notated score involved annotating the spectrogram as a MIDI-layer in the software *Sonic Visualiser* and then exporting the MIDI-file and audio to *Finale* notation software for score editing.

a.



b.

c.

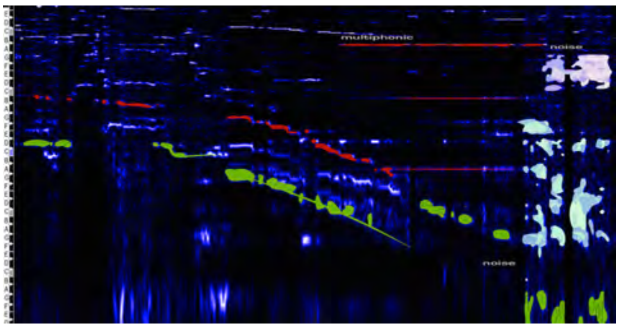


Figure 9. a. notated score, b. spectrogram score and c. annotated spectrogram score (excerpts) for *between states* [2015].

In the initial attempt I was unable to develop a satisfactory accommodation between the degree of detail in the score and the scroll-rate required to read it: if it was over-detailed it had to move too quickly to

[1] A play on the concept of the “no-input mixer”, a microphone is player in the bell of a bass clarinet and amplified just to the point of feedback. The feedback can then be initiated and “shaped” by closing or opening keys on the instrument without any other impulse from the performer.

accurately read (Fig. 9a). For the first performance a simplified spectrogram score (Fig. 9b) was used with a pitch-guide fixed at the left of the score. The pitch-guide replaced the single-line "playhead" that is used in most scrolling scores in the *Decibel Scoreplayer.* I had used the same approach to read Percy Grainger's *Free Music* in the Macintosh version of the score player (and this had later been coded into the *iOS* version by Aaron Wyatt). The pitch-guide had the advantage of providing more detail to the performers more "morphological" information about the sonic shapes and could also be scrolled more slowly than a conventional score. However in this version each player was allowed to choose which shape they would render with their instrument.

A final version of the score (Fig. 9c.) identified sounds to be performed by each performer (bass flute - red, bass clarinet - green) and contains text annotations and hue variations to represent different timbres. Since the parts are more defined, the full (unsimplified) spectrogram was re-added to the score giving the performers a greater indication of the context in which their sounds are heard.

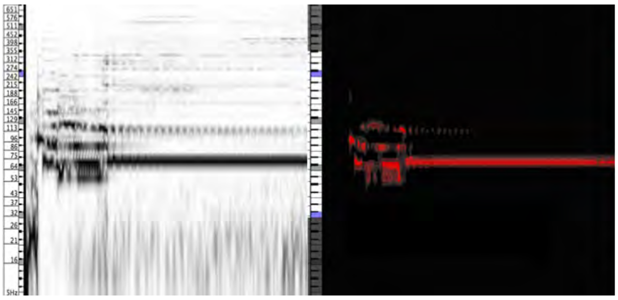


Figure 10. The opening of the *Ionisation* spectrogram (left) and the score of *residual drift* (right).

In *residual drift* the “representation sound to be produced by the musician” approach is explored to the fullest extent. The first few seconds of the *Ionisation* spectrogram and the score of residual drift are shown in Fig. 10 (note the "piano roll" style pitch indication is used as a "playhead" for the scrolling score to orientate the performer).

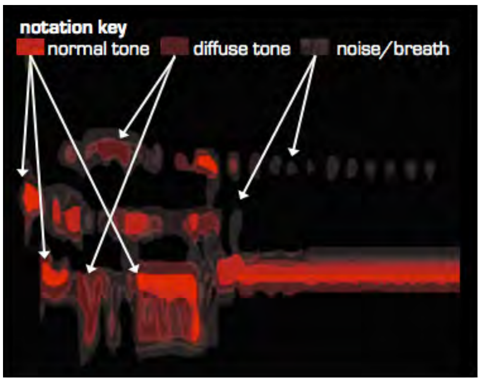


Figure 11. Notation key for *residual drift* [2015].

But the "piano roll playhead" is the only concession towards annotating the score: the performer is given a key classifying three hues with varied timbral qualities (Fig. 11) and then the performer is left to their own devices (along with the resynthesised spectrogram recording that plays synchronously with the score) to determine precisely how the sounds will be actuated.

The notation for *the miracle of the rose* [2015] combined procedures from the spectrogram score with gestural conventions. The work is based on a passage concerning the time-altering nature of solitary confinement from Jean Genet’s novel *The Miracle of the Rose* [1946]. In it Genet considers the mastery of time through the performance of gestures with infinite slowness - that "Eternity flows into the curve of a gesture". A collage of time-stretched recordings of the text by Australian/French artist Emmanuelle Zagoria – becomes the gesture that is curved in time. The spoken phrases are transcribed for the two percussionists into gestures exploring their cadence and timbre via varied instruments and notational approaches.

In Fig. 12 the notation for player 2 (which occupies the lower half of the page), indicates the amplitude of the sound (the vertical height), the timbral richness (hue), onset of event (stem) and direction of the bow (beam). The notation was created using software that I created for the project *Lyrebird: environment player* that was originally “intended to visualise sonic features of a ‘field recording’” (Vickery 2014). In this software the amplitude and frequency of the single strongest detected sinusoidal peak is represented by the size and vertical height of the rectangles drawn on a scrolling LCD object (in this case *jit.lcd*) and, brightness, noisiness and bark scale data derived using Tristan Jehan’s `analyzer~` object are used to determine the luminance, hue and saturation of each rectangle. This allows for the Scoreplayer to visualise timbral features of a recorded sound.

In the upper half of Fig. 12 the notation for player 1 indicates muted cymbal strikes (speech rhythms transcribed from the spectrogram). The changing position of the strike is indicated by the direction of beam. In both parts the thin beams indicate the movements of the performer’s arms between actions. The score is intended to be projected behind the performers allowing the audience to see the ritualistic gestural coordination between the performers and the score.

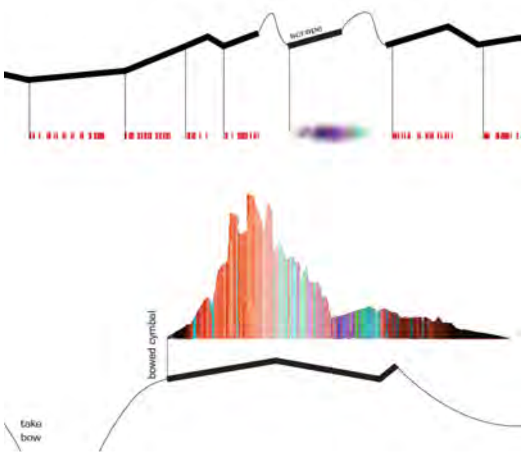


Figure 12. Excerpt from the score of *the miracle of the rose* [2015].

In Fig. 13. player 2 is bowing 5 different keys on the vibraphone, while player 1. is lowering a medium-sized chain onto the keys of a second vibraphone. The red note-heads and the downwardly inclined beams indicate lowering the chain onto the keys and orange note-heads (and upwardly inclined beams) indicate lifting the chain off the keys (both actions produce sound). Again the contours are transcribed directly from a sonogram of the accompanying recording.

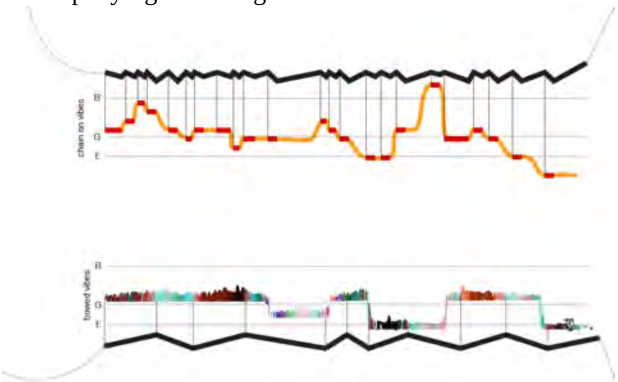


Figure 13. Excerpt from the score of *the miracle of the rose* [2015].

These works explore varied aspects of what O'Callaghan terms "mimetic instrumental resynthesis":

> Not only do these works use 'extra-musical' source materials as the starting point of their analyses, but they also attempt to preserve aspects of the source sound through the transcriptive process to engage in a mimetic discourse. (O'Callaghan 2015)

In particular the works allow for the contrast and interaction of instrumental and machine forms of sonification.

4. AUDIO PROCESSING FROM SOUND AND IMAGE

Many of the possibilities opened up by the processes described above have been enhanced by developments afforded by the *Decibel Scoreplayer* namely: synchronised networking, communication with external computers via OSC, audiofile playback, cross-fading of layers, random playback of score "tiles" and "nesting" of scoreplayer types (Hope et al 2015). The ability to interchange audio and visual representations of a work and to precisely synchronise them with live electronics provides a controllable work environment, including performers and electronic sound, that is not unlike that only previously available to Acousmatic composers working with recorded sound alone. The arrangement comes very close realising the "computer controlled performance environment" proposed in my 2011 paper "The possibilities of novel formal structures through computer controlled live performance".

The realtime transcription of field recordings into a symbolic score for an improviser to perform with or around in *Lyrebird: environment player* has been previously discussed (Vickery 2014). In *murmurs trapped beneath the bark* [2014], however, the source audio is a processed recording of a clarinet improvisation. Its ambiguous resemblance to real-world natural sound led me to term this an "artificial filed recording" and indeed the score produced by *Lyrebird* here has itself been visually processed in *Illustrator* to further the analogy with the recording (Fig. 14). This work, and this process opens up the possibility of creating hybrid real/mimetic sound works interchangeably combining field recordings and their machine and human emulations.

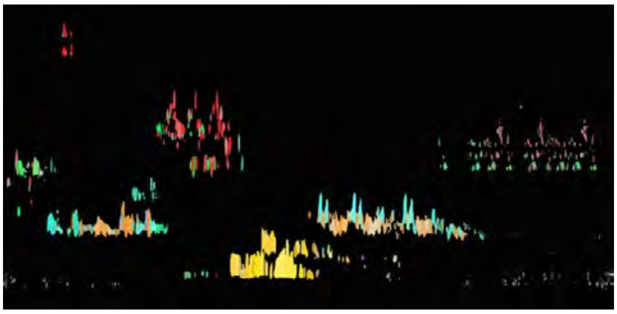


Figure 14. Detail from the score of *murmurs trapped beneath the bark* [2014],

Both *trash vortex* [2015] and *detritus* [2015], for example, feature scores comprising three layers: a graphical score, a rhizomatic path and an image collage (Fig. 15). The "path" layer periodically (and indeterminately) becomes transparent causing the graphical score to appear to "submerge" into the background collage - making it more challenging for the performers to read. The score (in the Decibel Scoreplayer) sends messages about the "path" layer's status via OSC to a *MaxMSP* audio processing patch, which in turn alters the audio processing of the live instruments to reflect the state of the score.

Figure 15. *trash vortex* [2015] three layered score: a graphical score, a rhizomatic path and an image collage.

My work in progress, *split constant* takes the opposite approach, repositioning score fragments in the Scoreplayer via OSC messages from the computer to align the performer with varied audio processing strategies.

In *in nomine tenebris* the contours of four of the instrumental scores are used to determine the spatialisation of the audio over 8 speakers (Fig. 16).

The data was transcribed by tracing the contours of the instrumental parts into four `function` objects. The data is then retrieved by inputting the position of the audiofile as reported by `snapshot~` (or score as reported by a *Tick* message received from the Scoreplayer over OSC).

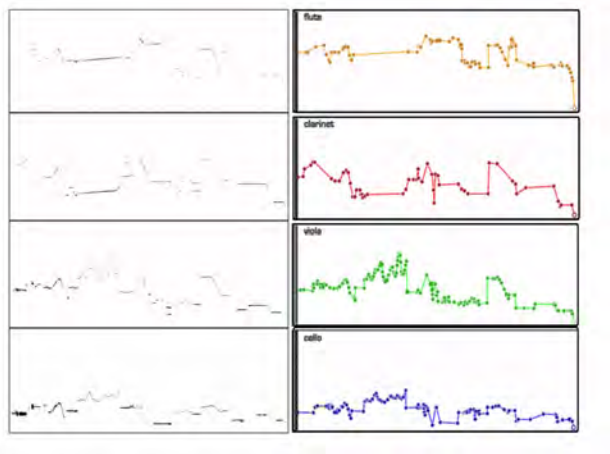


Figure 16. Contours of four of the instrumental parts for *in nomine tenebris* transcribed into four `function` objects.

The contour data also controls real-time spectral envelope warping of the acoustic instruments using Eric Lyons' `mindwarp~` in order to stretch instrumental spectra as the pitch of the instrument rises. *acid fury* in contrast, uses analysis of the found sound recording using data derived from Jehan's `analyzer~` object to spatialise the recording over 8 speakers in real-time. Similarly, *residual drift* employs real-time analysis of the noisiness of the live instrument to control the amplitude of the signal – allowing for inharmonic sounds to be favoured over harmonic ones.

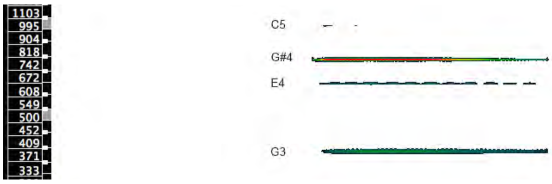
Similarly, *inhabited matter* [2015] and eight channel acousmatic work uses data derived from realtime analysis of its own structure to drive the spatialisation. The source audio in this case processed machine sonifications of the score of an earlier work *nature forms I.* This re-interrogating of materials is perhaps implicit in this project that emphasizes the interchangeability of audio, image and notation and, as noted, is also a feature of a number of other works discussed in this paper.

The score for *..with the fishes..* [2015] was built in a series of tableaux: oil rigs, flood, nuclear leak, deep sea, jellyfish/methane, submerged city, trash vortex. The tableaux were joined into a long image and then a score for viola, cello and double bass was then added on another layer. The score contained several references to musical objects: a recurring ship's bell and several passages from Debussy's *La Mer* [1905].

At an early stage of development of the work the instruments were recorded separately performing the notation and also performing passages with only the visual images from the tableaux. Once all of these elements had been created and had fixed temporal positions – the position of the scrolling score and the position of the recorded performances – they could be further compositionally developed through a range of interactions. The temporally fixed notation and audio allowed the process to proceed in a manner akin to manipulating audio in a Digital Audio Workstation: processing strategies could be auditioned; data could be derived from the score and used to control audio processing, pre-processed audio could be added and so forth.

At the simplest level the sound of a ship's bell was aligned with notation derived from a spectrogram of the same ships bell (Fig. 17) and short processed passages from Roger Désormière's classic 1950 recording of *La Mer* were aligned with the short quotations from the work.

a.



b.

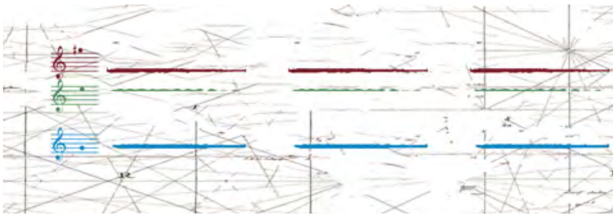


Figure 17. a. Spectrogram of ship's bell and b. notation of ship's bell in the score of *...with the fishes...*

Like *in nomine tenebris*, contour details from the background image were also mapped into `function` objects to control manipulation of the live performers using a range of processes including degrading the signal, pitchshift/delay and spectral manipulation (Fig. 18).

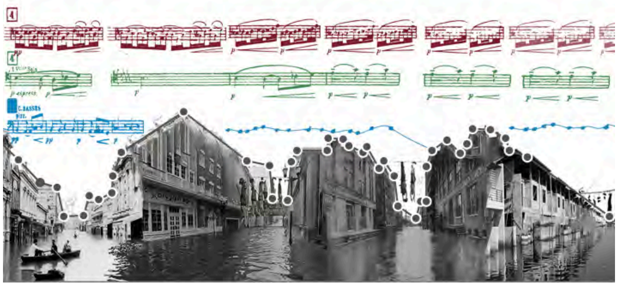
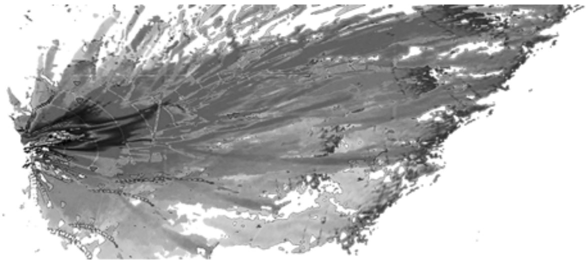


Figure 17. The *MaxMSP* function object mapped to contours of the score background image to control spectral manipulation of live performers and pitchbend spectrally "frozen" (with resent~) loop of Debussy's *La Mer* in *...with the fishes...*

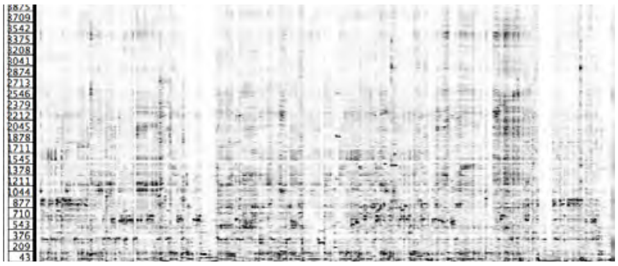
Audio files of the musician's performances of passages of visual images were then cross-processed against sonifications of the same images created in the *Sinereader* using Eric Lyons FFTease objects: namely resent, ether, thresh, shape, bthresher, pvwarp, pvgrain, disarray, taint, vacancy, burrow, codepend.

The processing included convolution with sonifications of the score image and the performance of the same image by the musicians to produce hybrid sounds combining both machine and performer realisations of the images. In Fig. 19 the original image (a.) is compared to a sonogram of instrumental reading of image a (b.) and the spectrogram of the convolution between sonifications of the image and its performance by the musicians (c.).

a.



b.



c.

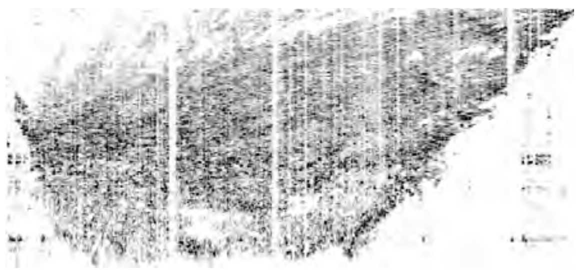


Figure 19. a. Image of Fukishima radioactive leak b. sonogram of instrumental reading of image a. and c. spectrogram of convolution of instrumental reading of the image and sonification of the image in *…with the fishes…*

5. CONCLUSION

These works explore the possibilities inherent in interchanging data between visual and sonic media. The processes involved including transcription of sound to image/notation, image/notation to sound, image to audio processing/spatialisation and audio to audio processing function interchangeably within a temporal framework. The imperfections in the transcription processes are intriguing in themselves as they throw into relief the distinctions between the various forms of representation. The implied circularity of processes opens the potential for re-interrogation of materials through continuous transmutation through the “eye of the needle” of transcription. These implications include the exploration of audio artifacts or “found sound” recordings as discussed but also extend to the exploration of field recordings and natural sound.

The efforts to extend notation discussed here are part of an ongoing project to better capture nuances of sound such as timbre, temperament and envelope morphology using shape and colour parameters (hue, saturation and luminosity). These experiments are further discussed in Vickery (2014a, b and c.)

Many of the processes described here are made possible through synchronisation of image, notation and audio afforded by computer networking, allow for the composer to operate simultaneously and interchangeably in both media. The inherent limitations of the analogy “that image and sound can be equivalent” can themselves be a rich field of investigation.

6. REFERENCES

Ablinger, P., 2011. Quadraturen 1995-2000. http://ablinger.mur.at/docu11.html

Cassidy, A. 2013. Constraint Schemata, Multi-axis Movement Modeling, and Unified, Multi-parametric Notation for Strings and Voices. Search Journal for New Music and Culture.

Barrett, D. G. 2007. Music and Its Others. http://ablinger.mur.at/docs/barrett_others.pdf

Bertin, J., 1967. Semiologie Graphique; Les Diagrammes - Les Réseaux; Les Cartes . Edition Gauthier –Villars.

Hope, C., Wyatt, A. and Vickery, L. 2015. The Decibel Scoreplayer: Enriched Scores For The iPad. Proceedings of the ICMC 2015, Denton, Texas 314-317.

Lachenmann, H. 1996. Musikals existentielle Erfahrung: Schriften 1966-1995 (Wiesbaden: Breitkopf & Hartel).

Morris, P. n.d. The Music Notation Project. http://musicnotation.org

O’Callaghan, J. 2012. Mediated Mimesis: Transcription as Processing. Proceedings of the Electroacoustic Music Studies Network Conference Meaning and Meaningfulness in Electroacoustic Music, Stockholm

O’Callaghan, J. 2015. Mimetic Instrumental Resynthesis. Organised Sound (20:2), 231 - 240

Vickery, L. R. 2011. The possibilities of novel formal structures through computer controlled live performance. Proceedings of the Australasian Computer Music Conference 2011, Auckland 112-117.

Vickery, L., 2014a, Notational Semantics in Music Visualisation and Notation. *Proceedings of the Australasian Computer Music Conference 2014*, 101-112.

Vickery, L. R. 2014b. Exploring a Visual/Sonic Representational Continuum. International Computer Music Conference. A. Georgaki et al. Athens, Greece. International Computer Music Association. 1: 177-184.

Vickery, L., 2014c. The Limitations of Representing Sound and Notation on Screen. *Organised Sound*, 19(3), 215-227, Cambridge, UK, DOI: 10.1017/S1355771814000193.

Wyatt, A., James, S., Vickery, L. and Hope, C. 2013. *Decibel ScorePlayer App*, Apple iTunes, https://itunes.apple.com/us/app/decibel-scoreplayer/id622591851?mt=8.

ACMC2015 – *MAKE!*

SUBMISSIONS BY ABSTRACT

LUDIC HUMAN COMPUTER CO-CREATION

Liam Bray *Oliver Bown*

ABSTRACT

Generative composition tools are now available on many mainstream computing platforms. They are more accessible and more usable than they ever have been before. Generative composition tools use autonomous processes as means of enabling a user's search for outcomes. The autonomous qualities of generative systems make them well suited to act as collaborative agents in a co-creative process. They do not necessarily exhibit the properties of human creative agents, but can stimulate and actively contribute to a creative process in a number of ways. Generative tools extend users' ability to innovate, promote discovery and can provide new ways of looking at familiar problems.

An initial study we conducted looked at utility and user experience. We compared how users develop strategies to complete both complex compositional tasks and open ended composition. The study looked at composition tools which adopt either a linear non-generative composition process or a non-linear generative one. The results suggested people become tightly engaged with generative processes. We propose to understand this in terms of a ludic interaction, focusing on play as a means of engagement which supports self motivated exploration.

In this presentation we expand on this notion of ludic interaction with creative systems, drawing both on the notion of ‘play’ in a musical creative sense, and the use of gameplay principles in productive software use, and in video game design. We look at the relationship between play and co-creativity in order to gain an understanding of how to evaluate success and how to design interactions in generative composition tools.

SOUND AS MULTIPLICITY IN EXPERIMENTAL MUSIC: LISTENING WITH THOREAU, CAGE AND SERRES

Dr Ben Byrne

ABSTRACT

A proverbial question asks, 'if a tree falls and no one is there to hear it, does it make a sound?' The question is interesting not for its possible answers – which clearly depend on the definitions of sound, hearing and subjectivity employed – but for the further questions it raises. How is it that a tree falling is considered to 'make' sound? And, more tellingly, why is it that the tree is thought to make 'a' sound distinct from that of the wood of which it is part?

Idealised archetypes of sound such as the voice, sine tone and high fidelity recording have had a marked influence on the study of sound, supporting an assumption that sound can be approached as a thing, treated as a discrete, reproducible commodity. However, experimental music has been greatly influenced by those who have gone into the woods and listened to sound as multiplicity – sound as dynamic, heterogeneous, contingent and discursive. 'Sound as Multiplicity in Experimental Music: Listening with Thoreau, Cage and Serres' traverses experimental music that approaches sound in this way, drawing on the experiences of Henry David Thoreau, the music of John Cage and the philosophy of Michel Serres as well as the capabilities of digital technologies. This is music that is itself multiple and must be heard rather than played. Music that complicates accepted notions of identity, mediation and environment, shifting emphasis from composers and performers to listeners.

ACOUSMATIC COMPOSITION IN A BLENDED WORLD

Luke Harrald
Elder Conservatorium of Music Faculty of
Arts, University of Adelaide

ABSTRACT

As we move into the era of ubiquitous computing and smart cities, unprecedented opportunities exist for composers to move their work out of the studio, concert hall or art gallery and into public space. While Augmented Reality is already upon us through various smartphone apps and other technologies, the general population will come to expect their real-world environment to become more responsive to their interactions through embedded technologies. Composers have an opportunity to engage in this through shaping the acoustic environment and offering new experiences for audiences that bend the relationship between composer, performer, and audience in unprecedented ways.

This paper explores recent site-specific installation works by the author, the shifting notion of diegesis within these works and their relationship with the audience. Particular attention is given to opportunities for the composer to engage with a site, approaches to blending the work with the existing acoustic environment, and approaches to moving the audience through different perspectives and states of immersion. Various computing and audio technologies have been used in the creation of the works, and these are examined from perspectives of computing, robustness, sustainability - or how to create digital art in environments without available electricity - and audio spatialisation in non-traditional environments.

The works have been built for a variety of locations for both temporary and permanent installation, including a remote abandoned mine, a steam train, parks, streetscapes, and a sports stadium. Lessons learned from site-specific work, are also discussed as to their re-appropriation into recent gallery installations by the author, in the context of the audience's understanding of place, and sense of Dasein or "being there".

GESTURE STUDY AND CORRESPONDENCES: COMPOSING FOR PERCUSSIVE IPAD APPS

Charles Martin
Research School of Computer Science
Australian National University
charles.martin@anu.edu.au

ABSTRACT

This paper describes the process of creating three compositions for multi-touch iPad apps and percussion: Gesture Study No. 1, Gesture Study No. 2, and Correspondences. These works were composed in 2014 as part of an effort to develop repertoire for two Canberra-based iPad groups, Ensemble Metatone, and the ANU New Music Ensemble. The works are for an ensemble of flexible size, and are open-ended with respect to the app that is used. Two of the compositions (Gesture Study No. 1 and 2) are for iPads alone, and one (Correspondences) is for a setup of iPads and small percussion instruments. We discuss how we have used standard musical notation to represent a vocabulary of percussive touch-gestures identified in previous research and how these works have adopted themes and techniques from percussion ensemble repertoire. Not only have these works led to several valuable performance experiences and an ongoing touch-screen ensemble practice, but they have also been used to explore the limits of our gesture-tracking agent software in real world situations.

3D PRINTING MICROTONAL FLUTES

Terumi Narushima
Faculty of Law, Humanities and the Arts
University of Wollongong

Kraig Grady
Faculty of Law, Humanities and the Arts University of Wollongong

Matthew Dabin
School of Electrical and Computer Engineering Royal Melbourne Institute of Technology

Christian Ritz
Faculty of Engineering and Information Sciences University of Wollongong

Stephen Beirne
Australian National Fabrication Facility
University of Wollongong

ABSTRACT

Microtonal tuning is a rapidly expanding area in which musicians are experimenting with a diverse range of tuning systems in their search for new resources for making music. One of the challenges in the field of microtonality is that conventional musical instruments are inadequate for realising the abundance of theoretical tunings that musicians wish to investigate. Already there is significant interest in instruments with microtonal capabilities such as keyboards (Narushima 2013) and re- fretted guitars (Schneider 2013) but very few microtonal wind instruments are available commercially. The aim of this project is to explore the potential for 3D printing to create custom-designed wind instruments that can play music in a variety of microtonal scales. Flutes serve as the initial stage in an investigation that can be extended to other woodwind and brass instruments in the future. We chose to focus on two types of flutes: the recorder, because it is relatively easy to play and maintain a stable pitch, and a simple transverse flute.

The project to date has involved designing and printing flutes based on pre-existing models, then extending this work to explore the effects of modifying several variables, such as the position and size of tone holes, as well as the shape and dimensions of the bore of the instruments. These are parameters that normally cannot be varied using standard manufacturing methods. We have also developed two different types of mouthpieces for a recorder and transverse flute, with the aim of producing a set of microtonal flute bodies with interchangeable mouthpieces. The long term goal of this new approach is to create a system in which, instead of the manufacturer dictating the tuning, customers are able to specify the tuning of their instrument for their own unique needs and have it printed on demand.

1. REFERENCES

Narushima, T, 2013. Mapping the Microtonal Spectrum Using Erv Wilson's Generalized Keyboard, (PhD thesis), University of Wollongong, Australia.

Schneider, J. 2004. "Just Guitar." Guitar International, 6: 42–50.

WRITING THE WRONGS OF HISTORY: LISTENING AND THE DISCOURSE OF GENDER IN ELECTROACOUSTIC MUSIC

Michelle Stead
Western Sydney University

ABSTRACT

This paper will discuss systems of value surrounding electroacoustic music by engaging with the epistemological formation of its discourse in relation to representations of gender. I will argue that history books have perpetuated and thus solidified androcentric versions of the history based purely on assumptions that women did not exist. Whilst these claims have, to some degree, been unpacked and scrutinised in the broader Western art music realm, they are still very much axiomatic to electroacoustic music discourse and literature. In more recent times crucial work has been undertaken to address this exclusion (see for example Elizabeth Hinkle-Turner and Tara Rodgers). However, this lack of female representation within academic literature, I argue, goes well beyond visibility. I consider the formation of knowledge about the music as gendered (especially in relation to the way the music is written and spoken about) and suggest this significantly impacts the way we listen to it. The aim of this paper then is to render visible the mechanics of exclusion thereby calling into question the foundations by which taken-for-granted knowledge becomes accepted as fact and, through my examination of the creative capacity of this gendered discourse, I will imagine how this might be taken up in listening practice.

ACMC2015 – *MAKE!*

ARTIST TALKS

GREEN SKIES…. MOMENTS

Roger Alsop
Performing Arts, VCAMCM
Melbourne University

Dr. Brigid Burke
Independent Composer/Performer/Artist
www.brigid.com.au

ABSTRACT

Green Skies…. Moments is a series of experimental 8 minute films for the Federation Square Big Screen. It will be created over a period of 10 months within a calendar year. It is designed to interpret the changes that occur in a public space over a year, focussing on changes in use, environment, and inhabitants, documented through film, as audio-visual works. Five outcomes are intended for each month: sonic/musical responses to the film, and an audio-visual interpretation of the film by each artist, and a collaborative interpretation that amalgamates the outcomes to create a third.

This paper discusses the individual and collaborative processes used by each artist to create the works. First it details the processes used in creating *Green Skies ... Moments*; then the individual approaches each artist takes to creating each work, and finally how the artist integrate those approaches in developing the final collaborative works. A short demonstration of works so far will be shown.

1 INTRODUCTION

Burke and Alsop have worked together on a number of interactive audio-visual works, and each has a strong history of individually creating such works. Burke's focus is on integrating musical ideas with a combination of different media, as an exploration of her artistic process. Her activities include: performance, composition, improvisation, installation, collaborations with dancers, acoustic performers and other new media performers, print making, pen and ink drawings, painting and digital animation, and she brings these together under the umbrella of polymedia.

Alsop works across a variety of art forms, with a focus on mapping as a creative approach, interactivity, and cross-media in creating works in which all of the elements in use or expressed are interdependent. He too has worked in dance, installation art, algorithmic composition, and collaborative art, and has also worked with communities in creating artworks and in documenting aspects of different communities.

The provocation behind *Green Skies ... Moments* was an invitation to Bourke to create a series of works for the Federation Square Big Screen, she then invited Alsop to be part of the developmental process and it was decided that taking a multifaceted approach would best represent each artist and their particular approach. The decision to create a final collaborative work each month that results from the individual works by each artist indicates the collaborative, community oriented nature of the Square itself, and this is what will be handed to Federation Square to be presented on the big screen.

2 INTENTIONS

The process for developing *Green Skies ... Moments* has five steps. The first step is for Alsop and Burke to film the environment of Federation Square. Here the focus will be on natural events. The intention is to represent the effect of how changes in the natural elements will create a visual score and provide a sonic narrative over the time in which *Green Skies ... Moments* is being created. The artists will then select from the films taken to create the various works. These need not be the same films, and it is possible for each artist to use a different film as a visual score and as the basis for inspiration for the audio-visual works developed from the selected film. The final work will be developed collaboratively from the resulting individual works, and the parameters of the collaboration will be decided on a work-by-work basis.

The visual images will be used to create a score which may be interpreted through any sound making device, this may be traditional instruments, instruments created from natural objects, or instruments created within a computer. The sounds captured in the film may be used as source material, as inspiration for a piece, or as indicators for different approaches to sonic composition. By taking these different approaches diverse works will be created from Burke and Alsop's different interpretations. These differing interpretations will create diverse outcomes based on the same events, and these outcomes will be amalgamated in the final collaborative work.

While Federation Square is in effect designed as a communal meeting place for people to inhabit, and therefore people are its focus, *Green Skies ... Moments* considers that the natural elements have a major effect on the public's use of that space and the behaviour of people within it. By documenting and interpreting these changes, it is anticipated that a interpretive diary of Federation Square will be created.

3 TWO CREATIVE APPROACHES TO GREEN SKIES

Each author gives the following two descriptions of the creative processes in their own words. They demonstrate the particular idiosyncrasies each take to the creative process, and how very different works can be derived from the same images and images taken form the same location.

3.1 Brigid Burke

The first thing that I look for is an image of Federation Square that for some reason strikes me as charged with meaning, even if I cannot formulate this meaning in discursive or conceptual terms. See two raw images of Federation Square in Figure 1 and 2.

Figure 1 Federation Square blurred raw image

Figure 2 Federation Square raw image on a clear day

As soon as this meaning becomes sufficiently clear in my mind, I set about developing it into a story, or better yet, it is the images themselves that develop their own implicit potentialities, the story they carry within them. Into the organization of this material, which is no longer purely visual but also conceptual, their now enters my deliberate intent to give order and sense to the development of the story, see figures 3 and 4. The images have been processed with other abstract paintings and drawings I have created. By layering these drawings they evoke a scene aimed to create themes in the musical composition through the pulse of film clips and the still images density of colours.

Figure 3 Federation Square processed

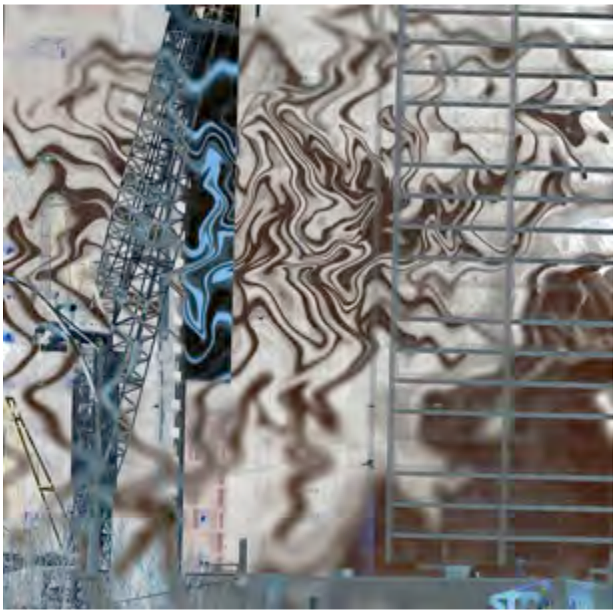


Figure 4 Federation Square processed

My approach is intuitive and evolutionary, and unique to each situation. The footage in these two photographs were both taken on cold gusty days these two situations created a anguished and fast pulsed movement in the music which I will interpret intuitively in performance and when creating the anosmatic sound.

3.1.1 Live Performance

The samples are made into a live interactive performance work using clarinet gestures, and electronic sounds from Federation Square to trigger visual samples.

The music component by Brigid is a combination of clarinet gestures depicting wind, fragmented chattering that are processed with granulations, reverbs, room placements and distortion and incorporates effects that warp and laminate and are mixed with synthetic samples of computerised wind and textured pulses of low frequencies samples. They are reminiscent of the behaviour of light reflecting and distorting the patterns of light and wave shapes. An example of these abstractions can be seen in Figure 3 and 4.

The visuals of the buildings and shapes of movement and light move in subtle ways through a spectrum of colours to depict the environmental aspects first envisaged for this composition, with added components different environments throughout the composition. A series of blue images, blue stills in pen and ink, silkscreen depicts the reflections created by light as seen in these images. Examples of these pen and ink drawings are seen in images in figures 3 & 4

3.2 Roger Alsop

Alsop's approach to interpreting the films is to see them as simply integrated visual and aural elements, without any focus on narrative or in relation to the location. He then examines them aesthetically as both documents of the events at the time of filming and as abstract aesthetic constructs. Figure 5 to Figure 7 are taken from one film shot in the evening of August 11, 2015. The film lasts for approximately sixteen minutes and the images are from the beginning, middle and end of the film.

Images from audio-visual works developed for *Green Skies ...* can be seen in Figure 8: % blending Figure 1 and Figure 5, Figure 9 <= blend of Figure 1 and Figure 2, and Figure 10 ~ blend of Figure 5 and Figure 6.

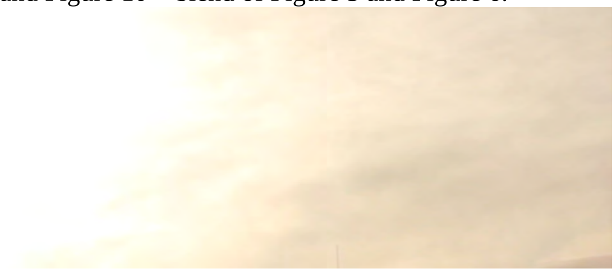


Figure 5: Light Sky

Figure 6: Light Sky with Bird

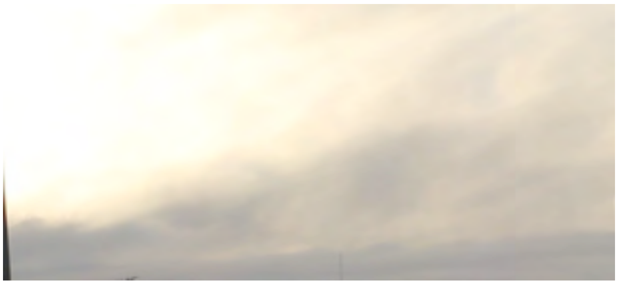


Figure 7 Darkening Sky

Figure 8: % blending Figure 1 and Figure 5

Figure 9 <= blend of Figure 1 and Figure 2

Figure 10 ~ blend of Figure 5 and Figure 6

These images are then represented in a 48 by 48 block matrix where the red, green and blue values of each block generate an audio signal using triangle, rectangle and saw waves, and the alpha value sets the amplitudes. The results form an audio response to the images as they are being generated and effected through the processes indicated.

Alsop has used the process of using pixel qualities to generate audio responses previously (Alsop 2007), where the mapped interaction between the image and the audio is used to create a sound work built from novel sonic interactions that the author may not have initially developed or considered. Like Burke he then allows for personal resonances within these interactions to emanate, and filters the outcomes to represent his aesthetic response, interpretations, and intentions. However there is no intention of "developing a story", or "exploring implicit potentialities".

4 CONCLUSION

While the outcomes may initially seem to fit the area of 'visual music', the approaches taken fit the range of those used by 'visual music artists', (DeWitt 1987, Dannenberg 2005, Evans 2005, Jones and Nevile 2005) where an intuitive interpretation is made and implemented, often through the paradigms of musical composition, and approaches to data interpretation (Lodha, Beahan et al. 1997, de Campo 2007, Park, Kim et al. 2010, Alsop and Giles 2015) are equally used. By Burke and Alsop first individually developing works that lean towards these two poles, by finally blending these approaches and works it is seen that new works, which demonstrate more than the sum of their parts, will be created.

Versions of the works developed will be presented at the ACMC2015 conference, where the processes used will be demonstrated.

5 REFERENCES

Alsop, P. R. 2007. Blue Angel. Loop Bar and Gallery.

Alsop, P. R. and V. Giles. 2015. Interpreting Data: Re-contextualizing data to develop approaches to musical composition. Korean Electro-Acoustic Music Society's Annual Conferenc, Seoul, Korean Electro-Acoustic Music Society

Dannenberg, R. B. 2005. "Interactive Visual Music: A Personal Perspective." Computer Music Journal **29**(4): 25-35.

de Campo, A. 2007. "Toward a data sonification design space map."

DeWitt, T. 1987. "Visual music: searching for an aesthetic." Leonardo **20**(2): 115-122.

Evans, B. 2005. "Foundations of a Visual Music." Computer Music Journal **29**(4): 11-24.

Galanter, P. 2009. Truth to Process – Evolutionary Art and the Aesthetics of Dynamism. Twelfth Generative Art Conference. Texas A&M University, Texas, MFA Department of Visualization.

Jones, R. and B. Nevile 2005. "Creating Visual Music in Jitter: Approaches and Techniques." Computer Music Journal **29**(4): 55-70.

Lodha, S. K., J. Beahan, T. Heppe, A. Joseph and B. Zane-Ulman 1997. "Muse: A musical data sonification toolkit."

Park, S. H., S. Kim, S. Lee and W. S. Yeo 2010. Composition with path: Musical sonification of geo-referenced data with online map interface, Ann Arbor, MI: MPublishing, University of Michigan Library.

Partch, H. 1949. Genesis of a Music: An Account of a creative work, its roots and its fulfillments. . London, Da Capo.

Schafer, R. M. 1994. Our Sonic Environment and The Soundscape of the Tuning of the World. Vermont, Destiny Books.

Weibel, P. and G. Jansen, Eds. 2006. Light Art from Artificial Light: Light as a Medium in 20th and 21st Century Ar. Ostfildern, Hatje Cantz Publishers.

IMITATION AND AUTONOMY AS A BASIS FOR MUSICAL AGENCY

Andrew R. Brown
Griffith University
andrew.r.brown@griffith.edu.au

ABSTRACT

Recent experiences of performing with interactive music systems have stimulated ideas for me about how my creativity is enhanced through these interactions and how much creativity might be attributed the computational processes. These ideas resonate with old philosophical questions around tool use and existential interactions with them (Heidegger 1977). They are also issues raised in more contemporary discussions around integrative music (Rowe 1993, 2001; Cope 1996, 2004). However, today's powerful computational tools and enhanced interest in machine learning and computational creativity bring these questions into sharp relief (Coulton 2008). In particular, for me, the following questions are pertinent. How creative is imitation (i.e., mimesis and reflexivity) and does autonomy correlate with musicality, or our perception of it? (Pachet 2006). How are these seemingly competing capabilities coming together in musical agency? (Pickering 1995). In this presentation, I will reflect on these issues with reference to selected literature and through reflection on my own recent creative practice in developing and performing with variously autonomous interactive music systems that include deliberate use of imitation. In particular three of my recent works will be featured, *Unity In Diversity* (2013), *CIM-biosis* (2015), and *Ripples* (2015) – that latter which I hope to perform at ACMC 2015.

1. REFERENCES

Cope, David. 1996. *Experiments in Musical Intelligence*. Vol. 12. Madison, Wisconsin: A-R Editions. Cope, David. 2004. "A Musical Learning Algorithm." *Computer Music Journal* 28 (3): 12–27.

Colton, Simon. 2008. "Creativity Versus the Perception of Creativity in Computational Systems." In *AAAI Sprint Symposium: Creative Intelligent Systems*, 14–20. AAAI.

Heidegger, Martin. 1977. *The Question Concerning Technology and Other Essays*. New York: Harper & Row. Pachet, François. 2006. "Enhancing Individual Creativity with Interactive Musical Reflexive Systems." In *Musical Creativity: Multidisciplinary Research in Theory and Practice*, edited by Irène Deliège and Geraint A. Wiggins,

359. New York: Psychology Press.

Pickering, Andrew. 1995. *The Mangle of Practice: Time, Agency and Practice*. Chicago: The University of Chicago Press.

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

Rowe, Robert. 2001. *Machine Musicianship*. Cambridge, MA: The MIT Press. Unity in Diversity (2013): https://youtu.be/Gmunxu3eK8g

CIM-biosis (2015): https://youtu.be/sUtRCiaqpn0

Ripples (2015): https://youtu.be/9AxslQWZ4-U

2. ARTIST BIOGRAPHY

Andrew R. Brown is Professor of Digital Arts at Griffith University in Brisbane, Australia. He is an active computer musician and computational artist. His research interests include digital creativity, computational aesthetics, music education and the philosophy of technology. He pursues a creative practice in computer-assisted music performance and audio-visual installations, with focus on generative processes and algorithmic music performance including live coding.

"A PLETHORA OF POLYS" – A LIVE ALGORITHMIC MICROTONAL IMPROVISATIONAL COMPOSITION FOR IPAD

Warren Burt
Box Hill Institute
Institute of Higher Education – Department of Music

ABSTRACT

"A Plethora of Polys" is a polytimbral, polyrhythmic, poly-microtonal algorithmic live performance composition for a collection of iPad apps which uses various capabilities of the software in unique and unusual ways. Using the interconnectivity of AudioBus and CoreMIDI, microtonally enabled apps on the iPad can control each other's tuning in both expected and unexpected ways.

1. INTRODUCTION

"A Plethora of Polys" is a polytimbral, polyrhythmic, poly-microtonal algorithmic live performance composition which uses the interconnectivity of CoreMIDI, microtonally enabled iPad to control each other's tuning in both expected and unexpected ways. By combining the tuning possibilities of several programs, as well as getting desired microtonal scale, we can also get unpredicted microtonal resources based on mistakes in one program controlling another.

The principal apps used in the piece are Thumbjam, by Sonosaurus LLC (Sonosaurus 2014), Gestrument by Jesper Nordin and Jonatan Liljedahl (Nordin 2015), Jasuto Modular by Chris and Amanda Wolfe (Wolfe and Wolfe 2013), BirdStepper by Travis Henspeter and Beau Jeffrey (Henspeter and Jeffrey 2014), and Audiobus, by the Audiobus Team (Audiobus 2015). Additionally tuning files for the program were made on Wilsonic, by Marcus Hobbs (Hobbs 2015), and ScaleGen, by Jesper Nordin and Jonatan Liljedahl (Nordin 2015). Given the interconnectivity being developed in the iOS environment, clumsy though it is (with the use of iTunes still being necessary to transfer some data between programs), one can now begin thinking of a collection of apps as a modular environment, a set of composing potentials similar to the patching-thought of older analogue synthesizers.

2. THE TUNINGS

Wilsonic is an app which allows one to explore a portion of the tuning universe opened up by Ervin Wilson (Wilson 1969). There are a number of tuning formations developed by Wilson which assemble scale complexes by multiplying various harmonics against each other, in different combinations. The Hexany, for example, makes six-note scales by taking all the possible two-element products of a set of four harmonics. Various extensions of this idea then generate larger sets of pitches. Some of these are the Stellated Hexany, the Tetradic Diamond, and the Hexany Diamond. For this piece, I used the Hexany Diamond, which consists of all the ratios possible between 4 elements, plus the six possible ratios of all the 2 element products of the 4 elements. This sounds complicated, but immediately becomes clear when you see the diagram. The scales made with this formation have 19 tones, some spaced very closely together, functioning more as beating variants on each other rather than as single tones. Using this pattern, I made five scales, each of which uses successive combinations of consecutive odd numbered harmonics as their basis. That is, 5 7 9 11; 7 9 11 13; 9 11 13 15; 11 13 15 17; and 13 15 17 19; are the base set of harmonics for each successive scale. These scales are used by the program Thumbjam, and can be freely accessed at any time.

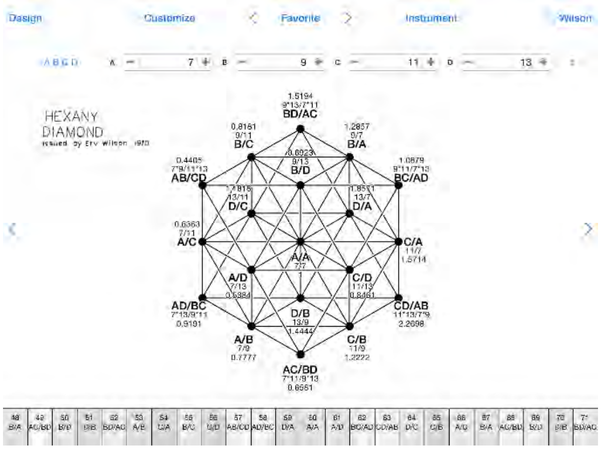


Figure 1 – Wilsonian Hexany Diamond in the Wilsonic app

Here is a listing of the pitches in the 5 7 9 11 Heaxany Diamond scale.

0: 1/1 0.000000 unison, perfect prime
1: 11/10 165.004228 4/5-tone, Ptolemy's second
2: 10/9 182.403712 minor whole tone
3: 63/55 235.104252
4: 90/77 270.079867
5: 11/9 347.407941 undecimal neutral third
6: 14/11 417.507964 undecimal diminished fourth or major third

 ISSN 1448-7780

7: 9/7 435.084095 septimal major third, BP third
8: 7/5 582.512193 septimal or Huygens' tritone, BP fourth
9: 140/99 599.911676 quasi-equal tritone
10: 99/70 600.088324 2nd quasi-equal tritone
11: 10/7 617.487807 Euler's tritone
12: 14/9 764.915905 septimal minor sixth
13: 11/7 782.492036 undecimal augmented fifth
14: 18/11 852.592059 undecimal neutral sixth
15: 77/45 929.920133
16: 110/63 964.895748
17: 9/5 1017.596288 just minor seventh, BP seventh
18: 20/11 1034.995772 large minor seventh
19: 2/1 1200.000000 octave

ScaleGen is an app by Jesper Nordin and Jonatan Liljedahl, which was designed as an adjunct to their Gestrument, an algorithmic composition and performance environment. Many different kind of scales can be generated by ScaleGen, and these can be exported to Gestrument, or as Scala files, or as text files. As well, one can use ScaleGen as a performance environment to hear what these scales sound like. I decided to concentrate on sub-harmonic scales in Gestrument. The classic subharmonic scale starts with a very high frequency and then divides that frequency by successive integers, ie 2, 3, 4, 5, etc. This produces a scale which is the inverse of the harmonic series, starting with a descending octave, then successive fifth, fourth, major third, minor third, neutral third etc. descending rapidly to groups of successive microtones, each slightly smaller than the previous ones. Gestrument allows you to use any division factor to make a “subharmonic” scale. (And any multiplication factor to make a “harmonic” scale as well.) For my purposes, I decided to make six subharmonic scales based on division factors of .23, .29, .31, .37, .41 and .43, which are also successive primes. These produce scales which have a smaller interval at the top, and then get into closely spaced microtonal intervals much more quickly. They all sound similar, but each one has a slightly different harmonic character and a different starting interval. Here is a listing of the starting intervals of the scales.

Division factor	Starting Interval in Cents
.43	619
.41	595
.37	545
.31	467
.29	441
.23	358

In this graphic you can see the nature of one subharmonic scale with the larger intervals starting off at a high pitch, and the intervals getting successively smaller as you go down in pitch.

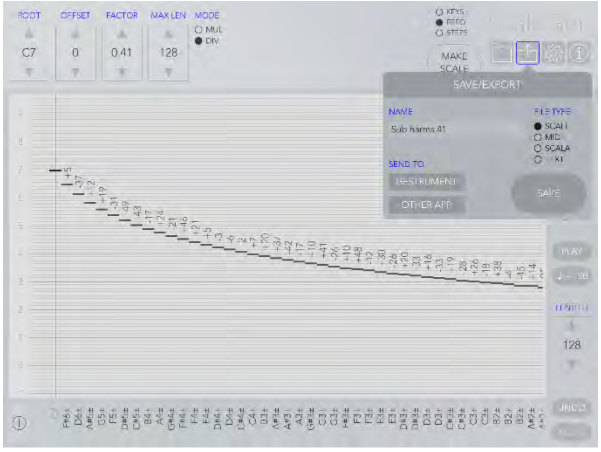


Figure 2. Subharmonic scale in ScaleGen app.

3. TUNING INTERACTIONS BETWEEN APPS

Both Thumbjam and Gestrument allow for microtonal scales to be used. They also both use the same algorithm for making microtonality, which is the “MIDI note number + MIDI pitch bend” routine. That is for each note in a microtonal scale, a table is made which has the MIDI note number and the MIDI pitch bend number needed to detune the note to the desired pitch level. This requires that each voice be monophonic, on a separate MIDI channel, but both allow for multi-MIDI-channel operation. For tuning accuracy, it also requires that the pitch bend range of the target synthesizer be set to +/- 2 semitones on each channel. This means that if Gestrument is used as a normally tuned MIDI generator, it can control a synthesizer that is set microtonally, and that synthesizer will then play in the desired scale. It also means that if Gestrument is set to perform microtonally, it can send “MIDI note number + MIDI pitch bend” sets to any desired target synth which is capable of “round-robin” channel assignment and has a settable pitch bend range. So, for example, if Thumbjam is set to normal tuning with a pitch bend range of +/-2 semitones, (with a starting MIDI channel of 1 and a MIDI channel limit of 16), it can play in whatever microtonal scale Gestrument is set to. And, as stated before, it also means that if Gestrument is set to play normal chromatic tuning, but Thumbjam is set to play a microtonal scale, Gestrument can control Thumbjam playing in its target scale.

Of course, that inevitably leads to the question, what happens if Gestrument is set to play in one microtonal scale, and Thumbjam is set to perform in another? The result here is a strange kind of hybrid scale, where the already detuned notes of Thumbjam are then modified, or retuned, by the MIDI pitch bend instructions of Gestrument.

Furthermore, Gestrument can play its own internal sounds (which can be any desired SoundFont .sf2 sound set), so these can be set to play in one scale, while controlling Thumbjam in another scale. This leads to lots of possibilities for strange doubling of pitches and timbres, which I exploit extensively in this piece.

Here are six sound examples: (These will be played live during the paper delivery from the iPad.)

1. Gestrument set to play chromatically; Thumbjam set to play Archytas' Enharmonic Genus, Dorian Mode. The result is a scale in the Enharmonic Genus.
2. Gestrument set to play Archytas' Enharmonic Genus, Dorian Mode; Thumbjam set to play chromatically. The result, again, is a scale in the Enharmonic Genus.
3. Gestrument set to play chromatically; Thumbjam set to play chromatically. The result is a normal chromatic scale.
4. Gestrument set to play Archytas' Enharmonic Genus, Dorian Mode; Thumbjam set to play Archytas' Enharmonic Genus, Dorian Mode. The result is a strange hybrid scale. I could figure out what the pitches are, but for the moment I prefer just to be charmed by it, and use it for its found-object possibilities.
5. Gestrument and Thumbjam set as above, but with Gestrument's internal sounds turned on. A similar timbre is used to Thumbjam – a plucked string. Now we hear Gestrument's string playing normal Enharmonic, but Thumbjam playing in the hybrid scale.
6. Gestrument set to play Chromatically with internal sounds, Thumbjam set to play Enharmonically. Now we have Chromatic scale and Enharmonic mode juxtaposed.

There are other combinations that can be explored here, but you get the idea from these. With five different scales on Thumbjam, and five different scales on Gestrument (plus the chromatic scale on both), you can see that there are lots of combinations of scales to explore in this algorithmic composing/performing environment. And as is probably obvious, my aim here is not to explore any one scale thoroughly, but to have a wide variety of harmonic resources that I can rapidly move between. The historical influence here is more, say, John Cage and Lejaren Hiller's HPSCHD (Cage and Hiller, 1969), with individual tapes in each equal temperament from 5 to 56 tones per octave, rather than say, any work of Harry Partch's, which thoroughly explored the resources of one portion of a just intonation scale-complex.

4. GESTRUMENT AND ITS PERFORMANCE INTERFACE

Gestrument has an interesting performance interface. It consists of a grid. Across the top of the grid are a series of musical durations, which are selectable by the user. Pitch is given by vertical position, duration by horizontal. There are also four sliders which affect the performance. Top left is a "Pulse Density" slider, which determines what percentage of the time a note will be played in the given rhythmic setting. Top right is a "Scale Morph A>B" slider – one can "morph" between two different tunings with this slider. The tunings are determined by Scala files, or ScaleGen files. On the left border, at the top is a "Pitch Fluctuation" slider. If this is set to 0, then a given part (of 8 possible parts) will only keep repeating the note the grid is indicating. At full on, the program will randomly select from a range of random pitches above and below the note the grid is set to. The range of this is set on the secondary page, in which settings and ranges of parameters for each of the 8 parts are given. Finally, on the left border, on the bottom, is a "Rhythm Randomness" slider. At 0, the program just produces notes at the horizontally given duration. At full on, it randomly displaces these durations by a quantization amount set on the secondary page. Combining this slider with the "Pulse Density" slider can produce quite a variety of rhythms. As stated above, there can be up to 8 voices controlled by this interface, and each one can have its own ranges for all of the parameters. This makes a powerful and flexible way of controlling music, once one gets one's head around the kinds of control made possible by this interface. What's more, the vertical pitch duration does not necessarily have to be low-to-high, as is usually the case. In the case of the subharmonic scales, in fact, the traditional pitch range direction is reversed, with a low position on the interface producing high pitches, and a high position on the interface producing low pitches. If a subharmonic scale is set on the left of the scale morph slider, and the chromatic scale is set on the right of the scale morph slider, then positions between these two will produce weird hybrid scales that tend to cluster around the middle of the pitch range. Having resources like this gives me a very rich environment to improvise within.

Figure 3. Gestrument performing interface

5. THE COMPLETE PATCH AND PERFORMANCE STRATEGIES

The iPad environment has grown in complexity and sophistication over the past couple of years, to the point now where I can construct pretty complex patches in various programs and then link those programs together. The Jasuto modular synthesis environment, by Chris and Amanda Wolfe, is an iOS or Android modular patching synthesis environment with a unique interface (One of the delights of the touch screen environment, whether iOS or Android, is the variety of physical interfaces that are being developed.) In Jasuto, modules, symbolized by differently coloured balls, are interconnected by patch cords. But the distance between the modules acts as an amplitude control for the strength of the connection. And the balls can be automated to move about the screen. So in this way, one can get a wide variety of changing LFO-style modulations happening within a particular patch. For this piece, I only wanted to use Jasuto to construct a particular kind of delay patch, so I didn't use the "motion" capability of the program. In this patch, the input sound travels immediately (through an allpass filter) to both channels of the output. The sound is also delayed 10 seconds and appears in the left channel, and also is delayed 20 seconds to appear in the right channel. A simple use of the program, but just what I wanted for this piece.

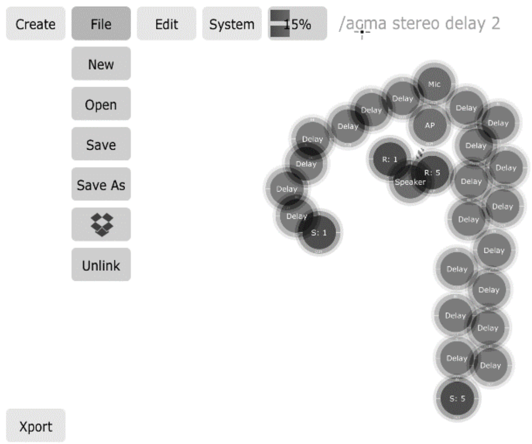


Figure 4. Jasuto GUI for this piece.

In this patch, the final effect unit is BirdStepper, an interesting set of 8 effects with an interesting control device. In this piece, I'm only using the "Spectral" effect in BirdStepper, but I'm controlling the 3 available parameters (time, feedback and gain) with hand drawn graphs which are stepped through very slowly. The "Spectral" effect produces a kind of "harmonic echo" on the input, which I found quite pretty and surprising in its sound.

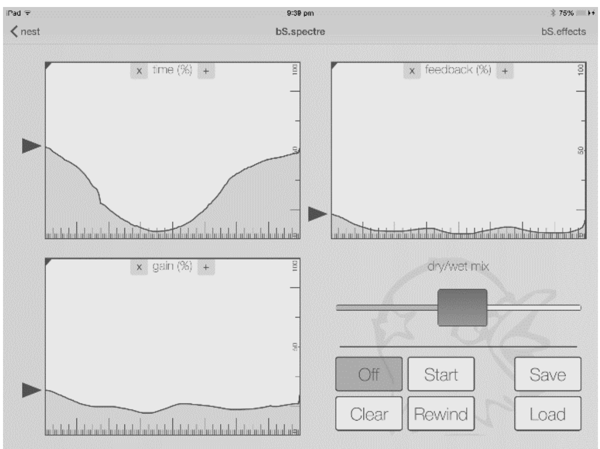


Figure 5. BirdStepper Interface

So with these four programs connected in Audiobus, I have a performance environment that is pretty powerful, and that I can improvise within. Here are some of the performing strategies that I can choose to use in an improvisatory performance using this setup.

Figure 6. Audiobus with complete patch

I could start with Gestrument's internal sounds turned off, but using the one of the Subharmonic scales to determine pitch. I could route its MIDI signal to Thumbjam, which I might have set to one of the Hexany Diamond scales. This will now produce a weirdly hybrid scale which I don't have complete control over. The output of this goes into the Jasuto delay and the BirdStepper "Spectral" effect, producing delays and harmonic arpeggiations of the gesture I performed with Gestrument. While performing this gesture, I can also change one of the four performance sliders in Gestrument, changing the nature of the gesture produced. I can also turn on and off a number of the voices in Gestrument, producing a texture of changing polyphony. I could then turn off the MIDI output of Gestrument, and turn on the internal sounds, thus producing Gestrument's sounds controlled in a "proper" version of one of the subharmonic scales. Again this will be processed by the delay and "Spectral" effect. Again, I can change the position of the performance sliders, affecting the textures I'm getting. As well, using the controls in Audiobus, I can turn off both Jasuto and BirdStepper, allowing the raw unprocessed sound from the synthesizers to be heard. I can also go into Thumbjam, and change both the patch and the tuning available in that app. As you can see, there are a lot of possibilities for performing here, and for getting different combinations of tunings and timbres, and different families of gestures in the piece.

6. FINAL THOUGHTS

The question might be asked, "Why do I want to compose a piece like this?" One answer might be that I find the combination of spontaneity and complex sounds produced by this patch to be very appealing, and I am delighted in the potentialities of these apps and their combinations. Although the “teen appeal index” of the techniques used in this piece might be low, the ornateness of the sound complexes produced by this patch are quite satisfying to my ears. Or, putting it more simply, these are sounds I want to hear, and if I don’t explore them, probably no one else will.

So using the iPad, I’ve created an algorithmic performing environment of great flexibility. The task now is to spend many hours performing this patch, going back and forth between the apps until I can do so with great ease and flexibility. In the time between when this paper is written (early October) and when the piece is performed at the conference (mid-November), I will have hopefully developed the required flexibility in performing so that a complex and engaging performance which alternates freely between tunings, timbres, varieties of melodic textures, and changing thicknesses of contrapuntal effects can take place.

7. REFERENCES

Audiobus Team 2015. Audiobus, iPad app, https://audiob.us/ (accessed Oct 10, 2015)

Cage, J. and Hiller, L. 1969. HPSCHD, Peters Edition, Henmar Press, New York

Cage, J. and Hiller, L., HPSCHD Available online at URL https://en.wikipedia.org/wiki/HPSCHD

Henspeter, T. and Jeffrey, B. 2014. BirdStepper, iPad app, Available online at URL http://www.birdsoundmusic.com/birdstepper/ (accessed Oct 10, 2015)

Hobbs, M. 2014. Wilsonic, iPad app, Available online at URL https://itunes.apple.com/us/app/wilsonic/id848852071?mt=8 (accessed Oct 10, 2015)

Nordin, J. 2015. Gestrument, iPad app, Available online at URL http://www.gestrument.com/ (accessed Oct 10, 2015)

Nordin, J. 2014. ScaleGen, iPad app, Available online at URL http://www.gestrument.com/scalegen/ (accessed Oct 10, 2015)

Sonosaurus 2014. Thumbjam, iPad app, Available online at URL http://thumbjam.com (accessed Oct 10, 2015)

Wilson, E. 1967. "Some Hexany and Hexany Diamond Lattices (and Blanks)" Available online at URL http://www.anaphoria.com/hex.PDF (accessed 10 Oct 2015)

Wolfe, C. and Wolfe, A. 2013. Jasuto Modular Synth, iPad app, Available online at URL www.jasuto.com (accessed Oct 10, 2015)

PLAY ONGAKU

Damian Castaldi
http://sodacake.com/

ABSTRACT

This artists talk describes the overall process in making a six channel audiovisual installation entitled *Play Ongaku*. It discusses the production process and examines how the installation invites an audience to engage with the work through playful, time based interaction.

音
楽

1. CONCEPT

This project is titled *Play Ongaku*, which is a play on words meaning *Play Music* in English and Japanese. The 'musical' element is significant and inspired by the Japanese symbol for Music 音楽 which is read ongaku meaning "music; musical movement". It is composed of the kanji 音 (read **on**) meaning "sound; noise" and 楽 (read **gaku**) meaning "music".

The 'play' element is reminiscent of Pierre Schaeffer and the 'Musique Concrète' movement within which the word jouer (to play) carried a double meaning: 'to enjoy oneself by interacting with one's surroundings', as well as to operate a musical instrument'.

The poster session and demonstration will address the theme MAKE! by discussing the overall process in making a six channel audiovisual installation and further discuss how the installation invites an audience to engage with the work through playful, time based interaction. Play Ongaku is centered around the assemblage of mixed media, acoustic, electronic and sampled material, the aesthetics of which have became the artist's domain. It identifies with his skill as a sound artist, drummer and percussionist, inventing new methods of playing found objects (such as bakelite phone bells and tin toy car wheels) and the sounds they produce using small mechanical hand made mallets inset onto mini servo motors and triggered by sensors within the cabinets.

Figure 1. Play Ongaku detail, October 2015.

Figure 2. Play Ongaku installation - Blue Mountains Cultural Centre Exposé exhibition Play for Time, October 2015

2. COMPONENTS

The electroacoustic cabinets trigger both digital sound design and amplified kinetic objects by close proximity using small Infra red sensors and a range of electronic and mixed media components including at its heart the Arduino Uno R3 microcontroller.

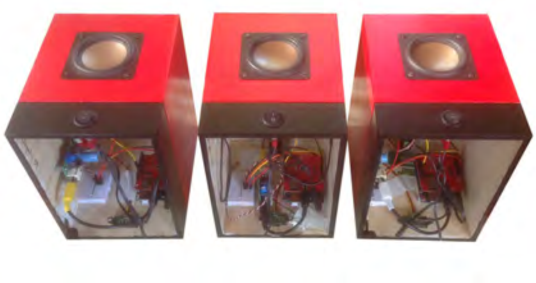


Figure 3. Play Ongaku detail - speaker side and internal electronic components, October 2015.

3. SIGNAL FLOW

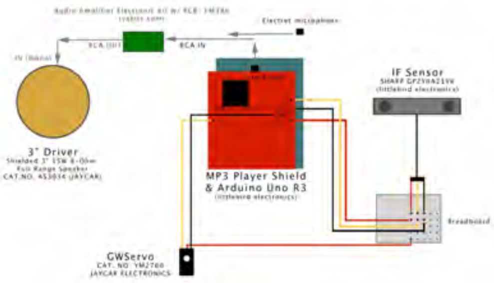


Figure 4. Play Ongaku signal flow for internal electronic components, October 2015.

4. GRAPHIC

The six images inserted on the top plate of each music cabinet were taken from the Japanese symbol for Music, which is read Ongaku 音楽. The symbol was spliced apart into 12 separate sections and graphically manipulated in Adobe Photoshop to form part of the whole when the six cabinets are placed together in two rows side by side. The images were printed on 300 gsm William Turner 100% cotton paper and glued with PVA onto 7mm plywood.

The six images are composed of the bottom section of kanji 音 (read on) meaning "sound; noise" and the top section of 楽 (read gaku) meaning "music".

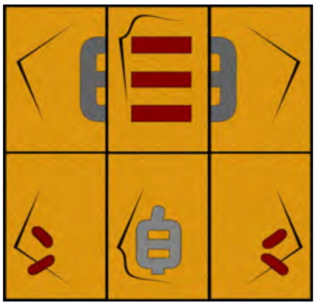


Figure 5. Play Ongaku graphic design - detail of six spliced and printed images, October 2015.

5. PROCESS

The skill set used to produce this work included sound design, digital audio production, location sound recording, woodworking, painting, soldering, wiring, electronic hardware design, graphic design, acrylic fabrication design, programming and coding. The cabinets are made with 4 sides and an inset top face for the printed image, pendulum and servo motor. A fabricated clear acrylic housing is fixed on top of each cabinet. The bottom of each cabinet is open to allow the power supplies to feed down through the plinth, which support each cabinet. The cabinets are screwed on top of the plinths on either side and are made from 7mm plywood. The dimensions are 211mm (D) X 149mm (W) X 190mm (H) based on a smaller version of the Yamaha NS10 studio monitor, which is 382mm x 215mm x 199mm. The cabinets are not sealed and contain one shielded 3" 15W 8-Ohm full range speaker.

6. AUDIO

The six channel audio mix is produced from individual mono audio files and live sound, which include:

- location sound recordings of a bird taken in my backyard, sampled and looped
- a bass synth rhythm, sampled and looped
- a recording of an old style mechanical clock, sampled and looped
- a recording of a scratch turntable and wheel, sampled and looped
- an Electret microphone pickup of a bell taken from an old style Bakelite phone
- an Electret microphone pickup of a tin wheel rolling over a metal trail, taken from an old style racing car found in a junk shop in Blackheath.

The digital audio was mixed, edited, processed and produced in the Logic Pro DAW and the export mono audio files were transferred to micro SD flash cards and inserted into the MP3 shields for audio playback.

Figure 6. Play Ongaku Logic Pro audio session, October 2015.

7. CREDITS

Play Ongaku is one of four installations produced in collaboration with composer/programmer Solange Kershaw for the Blue Mountains Cultural Centre Exposé exhibition Play for Time, exhibited in September/October 2015.

The artist would like to thank the Bundanon Trust for the valuable time spent in residence to produce this work, composer/programmer Solange Kershaw and the exhibition team at the BMCC for their help and support in exhibiting the work.

SONIC ART AS METAPHOR & MYTHOLOGY

Keith Halpin

Sound Design Lecturer, University of Gloucestershire
khalpin@glos.ac.uk

> A sound event is symbolic when it stirs in us emotions or thoughts beyond its mechanical sensations or signaling function, when it has a numinosity or reverberation that rings through the deeper recesses of the psyche.
> (Canadian composer R. Murray Schafer).

The author examines and explores the use of soundscape elements, mimetic sound- identities and spectromorphology to form metaphor and myth. A metaphor creates new links between otherwise distinct conceptual domains (Landy, 1991). Mythology is a vital feature of every culture. Myths are specific accounts in a time and place that is unspecified but which is understood as existing apart from ordinary human experience.

In our culture pictures have become tools used to elicit specific and planned emotional reactions in the people who see them. In calling up these deep emotions and memories, however, today's images have taken on new meanings and have created new myths that are shrouded-often deliberately-by these deeper memories. Sound also plays a pivotal part in shaping and recalling primitive emotions and memories.

It is proposed that electroacoustic music is a powerful medium capable of being used to arouse emotions, which manifest themselves in mental imagery, ultimately resulting in a formidable experience. Mythological themes in an unspecific time and place can be created, imagined and abstracted through electroacoustic music. The modern electronic composer has the power to guide this significative assignation through the interplay of a dominantly mimetic or a dominantly abstract sonic syntax. The listener is forced to re-evaluate their spectromorphological investigations in terms of a new reference point (Norris, M). This sort of covert and overt interplay can very powerfully affect the way a work is apprehended, and thus the structuring that the listener imposes. Therefore sound-images specifically spectromorphological (sound transformations) lead the listener to a deeper symbolic mythical reference point (Wishart, 1996). Moreover electrocacoustic music is a vehicle for the communication of mythical idea/narrative.

A large amount of Gestalt research is concerned with grouping and specifically to the way things are put together. Gestalt, meaning, “*form*", proposes that the whole is more
than the sum of its parts. Therefore the auditory system has its own version of "*perceptual complet*ion" (Levitin 2006). We are born with a predisposition toward interpreting sound, for instance if we hear a bird chirping outside the window our sensory perception forms a mental image of a bird. Composers can use this predisposition to skillfully influence their audience's emotions and generate mental images.

Daniel J Levitin (2006) states:

> Recording engineers have also learned to create special effects that tickle our brains by exploiting neural circuits that evolved to discern important features of our auditory system.

Goldberg (1992) explains that studies of the brain during a music therapy method known as guided imagery and music (GIM) present further insight into the connection between music, emotion and imagery. Goldberg argued that music "triggers emotions", which in turn "stimulates imagery".

The author will present and outline the symbolic relevance of combining two different sound-images to create new and thought-provoking meanings specific to mythology. The author will also discuss relevant composers and their postproduction techniques involved in constructing an electroacoustic composition for the communication of mythical idea/narrative.

Finally, several sound transformations will be played on loudspeakers to confirm the validity of his findings.

1. REFERENCES

Basanta, A. 2007. “An examination of the relationship between the aesthetic treatment of sound materials and the meaning conveyed by such means in text-sound composition”. Available online at www.sfu.ca/~aba36/

Davis, J. Power of Images: Creating the Myths of Our Time Available online at www.medialit.org/reading-room/power-images-creating-myths-our-time

Fagerlund, S. 2004. “Automatic Recognition of Bird Species by Their Sounds”, M. Sc. thesis, Helsinki University of Technology

Goldberg, F. 1992. “The Bonny Method of Guided Imagery”. in Wigram, T; Saperston, B; West, R. The Art and Science of Music Therapy. Amsterdam: Harwood Academic Publishers

Landy, L. 1991. “Sound Transformations in Electroacoustic Music”, pp.1-19, Available online at www.composersdesktop.com/landyeam.html

Levitin, D. 2006. This is Your Brain on Music. New York: Plume Books

Norris, Michael, n.d. An Overview and Assessment of Selected Discourse on Electroacoustic Music Available online at www.michaelnorris.info/soundmagic/diss2.html

Smalley, D, 1996, “The Listening Imagination” “Listening in the Electroacoustic Era, *Contemporary Music Review*. (13:2): 77-107

Wishart, T. 1994. Audible Design: A diagrammatic guide to sound compositional processes, York: Orpheus The Pantomine, pp. 1-74

Wishart, T. 1996. On Sonic Art. Amsterdam: Harwood Academic Publishers

2. ARTIST BIOGRAPHY

Keith Halpin is an Irish electroacoustic composer & sound designer working in acoustic, computer media, radio, TV & film.

Keith has an MA in Music Technology from the University of Limerick, BA in Media Production and HDip in Audio Engineering. He is an active member of the Audio Engineering Society and has extensive practice teaching undergraduates in a range of Music Technology disciplines.

Keith is a Senior Sound Design Lecturer for Gloucestershire University. He teaches Sound Aesthetics. This module progresses beyond basic approaches to sound in visual broadcast and looks in more depth at using sound as a form of creative expression and narrative structure. He is also the lead Sound Designer for King Sound Studios in Sydney, Australia.

His research interests include the use of soundscape elements, mimetic sound-identities and spectromorphology in composition to form metaphor and myth. Other interests include sound design for moving image and cloud based sound production applications.

ACMC2015 – *MAKE!*

STUDIO REPORTS

STUDIO REPORT: WEST AUSTRALIAN ACADEMY OF PERFORMING ARTS, COMPOSITION DEPARTMENT

Dr Cat Hope
Western Australian Academy of Performing Arts
Composition and Music Technology

Dr Lindsay Vickery
Western Australian Academy of Performing Arts
Composition and Music Technology

Jean-Michel Maujean
Western Australian Academy of Performing Arts
Composition and Music Technology

ABSTRACT

This studio report will discuss the Composition and Music Technology stream of the Bachelor of Music at the Western Australian Academy of Performing Arts at Edith Cowan University. Recent developments include: the reading and creation of digital graphic notations, image sonification, audio visualisation, online music documentation and archiving, rebuilding 21st century analogue electronics, investigating preservation and museological practises in the preservation of performing arts materials, wave terrain synthesis, timbral spatialisation, the philosophy of the digital audio workstation and innovations in the process of composition. This paper provides an overview of research undertaken within the studio, the results it has generated, and an overview of its dissemination.

IMPLEMENTING THE NEW BACHELOR OF SOUND PRODUCTION AT BOX HILL INSTITUTE

Timothy Opie
School of Higher Education (Music)
Box Hill Institute

1. ABSTRACT

In 2016 a new Bachelor of Sound Production degree will commence delivery at Box Hill Institute. This paper details the development and integration of this new degree within the existing framework at Box Hill Institute.

2. INTRODUCTION

In 1987 Ken Barker published the book *Identification of statewide training needs for the music industry report to Curriculum Services Office of the Technical and Education Board of Victoria* (Barker 1987). With this he justified and began running a music course at Box Hill College of Technical and Further Education. Over the next two decades the music program grew a strong local reputation. At this point the Institute rebranded itself as Box Hill Institute (BHI), and began delivering undergraduate degrees. The Bachelor of Applied Music was the first undergraduate degree to be delivered at Box Hill Institute. Currently BHI delivers 12 undergraduate Bachelor degrees, including an updated Bachelor of Applied Music, and two postgraduate Masters degrees, including the Masters of Music (Contemporary Practice). Commencing in 2016 they will now also deliver the Bachelor of Sound Production.

3. PROPOSING THE BACHELOR OF SOUND PRODUCTION

Over the past 12 years the Bachelor of Applied Music at Box Hill Institute has grown and expanded to the point where BHI has become Australia's largest specialized contemporary music program, and one of the largest in the world (NACTMUS 2014). The Performance stream in particular has been in high demand. Within the Bachelor of Applied Music however there was one stream with a slight identity crisis: Audio Production. On one hand it included a full music theory curriculum including aural skills and performance opportunities, and on the other it was aimed at audio engineers, music producers, and sound designers, and included engineering and other technical skills.

For some students this was a perfect blend, they got the musical knowledge they wanted, and also learned how to set up and run their own studios. For others however this dichotomy in musical perspectives could not be resolved. The skills and knowledge required for audio production and live sound mixing require significantly different approaches and techniques. They require the students to think of sound sonically and timbrally, rather than musically. Students need to be able to pick up on distortions, flutters, phase shifting, noise and signals, across multiple channels coming from multiple locations simultaneously, instinctually and fluently acting upon them with their mixing consoles and computers as if they were playing a musical instrument (Brown 2001).

With this dichotomy in mind it seemed the best option was to create a new degree that diverged from the Bachelor of Applied Music. In consultation with numerous audio engineers, live sound experts, and other industry leaders, a program was put together over a couple of years, with the final version being submitted and accepted by TEQSA in 2015. Incoming students for 2016 now have the option of enrolling in the existing Bachelor of Applied Music Audio Production stream, which shares units with the Core Applied Music Program, or they can specialise in a much broader field of sound production.

4. COMPONENTS OF THE NEW DEGREE

The structure of the Bachelor of Sound Production is intentionally very simple, belying a much more comprehensive and profound bachelor course. If you refer to Figure 1 on the following page you can see how these subjects fit together to create a cohesive new degree. The units are designed in a way that many projects can be integrated with particular Bachelor of Applied Music projects. For example the Performance students have projects that need to be recorded, which they will propose to the Sound Production students who will run the studio sessions. In order to convey an idea of what is covered in the course the following section broadly outline the contents of each subject area within the Bachelor of Sound Production

4.1. Studio Techniques

Studio Techniques covers all the techniques and knowledge required to work productively in a music studio environment, from simple microphone placement right through to multi-track mixing and post-production. It also includes certification in an industry standard digital audio workstation (DAW), which is currently the Pro Tools certification program, however this will be updated if another DAW takes its place in the industry. On top of the practical techniques required to master this area, Studio Techniques will cover aesthetics, production values and audio theory, including digital signal processing. By the end of Studio Techniques, every student will have created a commercial level, fully self-produced album, as well as many other smaller projects that can become part of a folio for future employment.

4.2. Live Sound

Live Sound covers the techniques and knowledge required to work in a venue for live performance or live broadcast. This covers areas from bands in pubs, to musical theatre, from Rockwiz to the tennis. It includes live mixing techniques, backline, sound checks, etc. Live Sound covers various communications protocols, especially concerning wireless equipment, and low latency networking for audio streaming. It also covers acoustics, including acoustic design, room analysis, and acoustic theory. Live Sound also covers aural, pertaining to the needs of an audio engineer. This includes tuning, keys, clipping, distortion, phasing, and many other audio issues that can be identified with a trained ear.

4.3. Music Production

Music Production covers a multitude of aspects of electronic music creation. This includes sequencing and arranging music, synthesis, sampling, and signal processing. Within the sequencing and arranging components, music theory will be covered in a practical manner, enabling students to create well structured musical works with effective harmonies, rhythm and melodies. It covers the history of music production technology, including the pioneers, practitioners and seminal works that define new advances in music production technology. Music Production also delves deeply into computer music programming, including MIDI manipulation, algorithmic composition, creating digital synths and plugins, using external hardware as controllers, and also hacking non-musical devices to use in live musical performance, composition, theatre, circus, and installation art works. Third year especially allows them to work on large-scale projects of their own design.

4.4. Sound For Media

Sound for Media focuses on sound design for television, film, documentaries, computer games, and other media. It includes the art of Foley, recording audio on set and on location, diegetic and non-diegetic music, and other sonic events. Sound for Media also examines SMPTE and linear musical environments, contrasting it with interactive non-linear environments. By the end of Sound for Media, all students will have a folio of works designed to help them seek work in film and computer games audio design.

BACHELOR OF SOUND PRODUCTION

I	6 Points 4 hours	6 point 4 Hours	6 point 4 Hours	3 point 2hours	3 point 2 hours
Semester 1 (1st year)	BSP111 Studio Techniques 1	BSP112 Live Sound 1	BSP113 Music Production 1	BSP114 Sound for Media 1	Elective 1
Semester 2 (1st year)	BSP121 Studio Techniques 2	BSP122 Live Sound 2	BSP123 Music Production 2	BSP124 Sound for Media 2	Elective 2
Semester 3 (2nd year)	BSP131 Studio Techniques 3	BSP132 Live Sound 3	BSP133 Music Production 3	BSP134 Sound for Media 3	Elective 3
Semester 4 (2nd year)	BSP141 Studio Techniques 4	BSP142 Live Sound 4	BSP143 Music Production 4	BSP144 Sound for Media 4	Elective 4
Semester 5 (3rd year)	BSP151 Studio Techniques 5	BSP155 Business Skills	BSP153 Music Production 5	BSP154 Sound for Media 5 (6 point)	
Semester 6 (3rd year)	BSP161 Studio Techniques 6	BSP165 Work Placement	BSP163 Music Production 6	BSP164 Sound for Media 6 (6 point)	

Figure 1. Overview of Units

4.5. Business Skills

Business Skills is a subject that prepares students to either seek work, or create new work opportunities. It is geared towards the music industry, and includes creating business plans, self-promotion, publication, project management, strategic and financial planning, business pitches and many other employment skills.

4.6. Work Placement

Work Placement is an opportunity for students to spend some time in a real work environment. In this capacity they will be expected to meet the needs of their employer, using the skills that have been covered throughout the rest of the course. Various opportunities will be presented, giving the students an option of trialing in an area that suits their field of expertise. The employers will provide feedback on their time there.

5. CONCLUSION

The Bachelor of Sound Production is really exciting. It has been a work in progress for a couple of years now, but a lot of consultation and thought was put into each specific area, producing a strong and cohesive whole. Of course the test will come when the students arrive, but we all feel it will prove successful.

6. REFERENCES

Barker, K. 1987. *Identification of statewide training needs for the music industry report to Curriculum Services Office of the Technical and Education Board of Victoria.* Box Hill, Australia.

Brown, A, 2001. How the computer assists composers: A survey of contemporary practice. In Knowles, Julian, Eds. *Proceedings Australasian Computer Music Conference*, pages pp. 9-16, University of Western Sydney.

NACTMUS, 2014. *NACTMUS Members, Victoria, Box Hill Institute.* Site: http://www.nactmus.org.au/nactmus-members/box-hill-tafe/ Accessed 16/10/2015

THE PERTH ARTIFACTORY: STUDIO REPORT

Meg Travers
Western Australian Academy of Performing Arts
Edith Cowan University
Perth, Western Australia, Australia

SKoT McDonald
Perth Artifactory
8/16 Guthrie Street, Osborne Park
Western Australia

ABSTRACT

The Perth Artifactory has become known not only as a ‘hacker space’, but also as a location for the making and performance of electronic and experimental music and instruments.

Keywords: electronic music, makers, hack, modular synthesizer, digital community, noise, studio report, Theremin, BollART, percussion, FXpansion, BFD.

1. HISTORY

The Perth Artifactory came into existence in 2009. It was originally located in a residential area in an inner suburb of Perth, but soon moved to a large warehouse in the industrial zone of Osborne Park where its members and the range of activities has grown.

The Perth Artifactory has always had a number of members with a strong interest in electronic and experimental music, and in the last three years this has become a substantial sub-group within the organisation, with activities, facilities and regular events that centre around this. There are also many people within this group with ties to the Western Australian Academy of Performing Arts, either as past or current undergraduate and postgraduate students.

2. ACTIVITIES

2.1. NoizeMaschin!!

Entering its fifth year, NoizeMaschin!! (The Perth Artifactory 2015) is a monthly experimental music night organised and curated by a team of people from the Artifactory’s music sub-group. The format for NoizeMaschin!! is unusual: each night includes 8 performers (these may be soloists or groups), and each act is given 10 minutes in which to perform, ideally segueing into the following performer. This short-set format means that many styles are heard over each evening, and allows for less “new music fatigue” in the audience, many of whom are newcomers to this genre.

All of the events have been recorded and made available for download (with the performers permission) and are being included in the Western Australian New Music Archive (WANMA).

Whilst the Artifactory is well known to many of the arts funding organisations in Western Australia, it does not receive any ongoing funding from them, and NoizeMaschin!! is produced on a zero dollar budget. The PA system and staging used for the event were found or donated by members who also donate their time to construct them and keep them running. Lighting for the events is often provided by one of the members who has constructed a computer controlled laser display, and also via a hybrid MIDI/DMX system where a controller keyboard is used to literally “play” the lights.

Performers understand that there is no capacity to pay them, however this hasn’t dented the enthusiasm of people wishing to perform and it has proved to be a useful avenue for many to trial experimental techniques before moving onto bigger shows.

NoizeMaschin!! also offers three month residencies to new performers which gives them free access to the Artifactory and its facilities, and the opportunity to perform at each NoizeMaschin!! concert over that period.

2.2. Modular Synthesizer building

Another of the regular events is the monthly Modular Synthesizer building night. Local synthesis guru and recent PhD graduate in electronic engineering, Dr Andrew Fitch, designs and builds Eurorack format synthesizer modules (nonlinearcircuits) which are sold worldwide. For these nights however, Andrew supplies the modules in kit form (including all of the components plus PCB board and panel) at very low prices, and assists people to build and test synthesizer modules, which they are literally building themselves.

Another of the members, Nathan Thompson, has designed and cuts inexpensive modular synthesizer cases on the laser cutters; these are also gaining popularity in the modular synthesizer community.

Usually 20-30 people attend these nights each month, with a high percentage of musicians and engineers. This has led to great cross-pollination between the two professions, and the growing possibility that Perth has the highest ratio of DIY modular synth owners in Australia.

2.3. The Sonic BollART

In 2015, the Perth International Arts Festival held the first Arthack in Australia – Hack the Festival (Perth International Arts Festival 2015). This event was a week long hack-a-thon in which ideas were discussed, teams formed, and art hacks created.

Figure 1. Dr Andrew Fitch and one of his nonlinearcircuits modular synthesizers.

The Perth Artifactory had several members on various teams in the event, but also fielded an official Artifactory team which created the Sonic BollART. The BollART was premised upon the idea that bollards are common objects in any urban landscape, serving as a barrier to certain activities or modes of transport, but that it also offers opportunities to house infrastructure for other things. Representing an underappreciated, readily available and numerous art-space with high foot-traffic, a human scale, and a fun set of constraints to fire the imagination, the BollART is a vessel within which a range of experiences can be tailored by individual artists, to create an item to convey their ideas.

The team’s entry, named the Sonic BollART, was constructed with the Artifactory’s laser cutter and contained a theremin and LED light show to inspire people to physically interact with the object. Whilst theremins are well known, most people have not had an opportunity to try playing one, and its hands-free performance interface provides no barrier to those wishing to attempt to play it.

Addressing the dual problems of distribution for unsigned musicians in the 21st century, and Perth’s current issues of finding audiences and venues for local music, The Sonic BollART also hosts a wireless, offline file-sharing service for local musicians to place their recordings, and for any member of the public to download these. In a nod to MOMA’s successful *Dead Drops* installation, this service provides a kind of Musical Parkour, giving a direct route for people in an area to discover music and musicians that are around them via a wireless data connection that does not rely on any other infrastructure, nor present a risk of over-utilising expensive 3G and 4G data services. Using the tagline “Un-cloud your music and find your local audience”, this system also encourages public discourse on the government monitoring and metadata collection from Internet-based systems.

The Sonic BollART was selected as the winner of the Hack the Festival event, and has since been exhibited at the Totally Huge New Music Festival, local rock music events, street festivals and TEDx.

Figure 2. The Sonic BollART

2.4. Audio and music Making

As a hacker-space, the Perth Artifactory has a range of advanced tools available for members to use, including laser cutters, 3D printers, CNC router and mill, lathe, welders, and so on, as well as a large space in which to design and create. Several musical instruments have been created there.

The Arcophone (Make Magazine 2011) is a MIDI controlled 30,000 volt electric plasma arc instrument which creates audibly tuned arcs of plasma from high voltage arc gaps (or Jacob’s Ladders). It is operated using a MIDI controller or software, a little trepidation, and a constant body-position-awareness of how close you are to the plasma arcs.

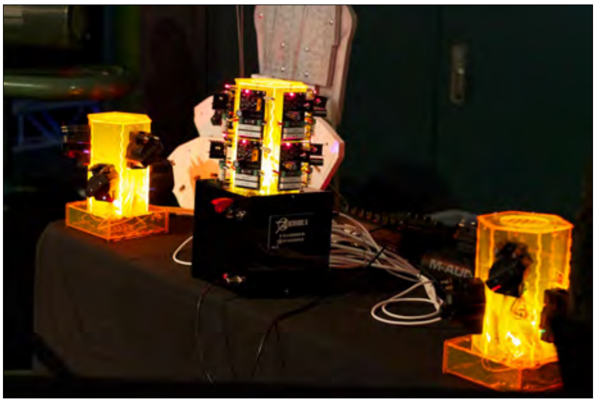


Figure 3. The Arcophone

With many people accessing the space but not necessarily being aware of how to correctly to power up the sound system, the Artifactory also has a network controlled audio system for the stage PA system. This system uses a Raspberry Pi to control power and the timing of switching on equipment via solid state relays. It includes an audio USB card to allow streaming of music from the network (eg: from a user's phone or laptop) and is operated only via a network connection, the hardware deliberately includes no switches or buttons for people to accidentally press or "explore".

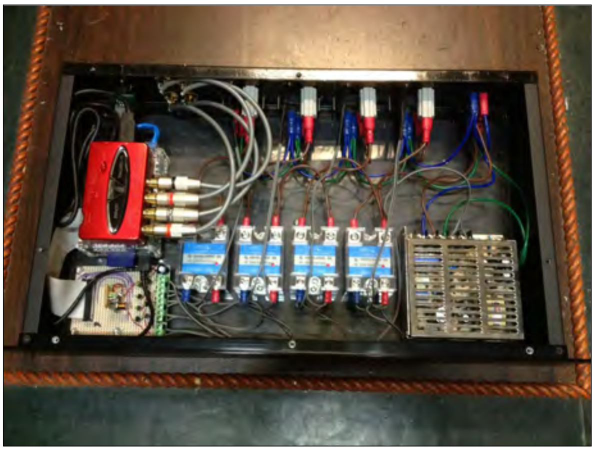


Figure 4. The sound system controller rig.

The Artifactory also hosts a rehearsal room which is operated under a membership system. Up to 5 groups share the rental of that portion of the warehouse giving them a secure, soundproofed room that has a limited number of tenants and the ability to leave large instruments in situ with minimal risk of damage or loss. This has not only assisted in paying the rent on the warehouse, but also provided further networking opportunities for electronic musicians who can promote and organize shared gigs, and to do small scale recordings.

The audio equipment in the room was donated or purchased using part of the proceeds from the Hack the Festival prize money. A local music store has also provided discounts to the users of the room and this has extended the networks of people in both organisations.

Figure 5. The rehearsal room, with obligatory lava lamp and Turkish rug.

Another monthly workshop is Arduino U at which members can be coached or be assisted with Arduino based projects. Following the International Computer Music Conference in Perth in 2013, a number of workshops specifically on Mozzi (Barrass) were held as a result of hearing Tim Barrass speak about this sound synthesis library which he created and maintains for the Arduino.

The construction of physical items is a big part of a hacker space, and with so much music and noise being generated both at the Artifactory and in people's homes, workshops have also been run on building sound dampening baffles. The CNC router *Swarf-o-mat* has also been pressed into use in the making of guitar bodies, and scavenged items have been constructed into a large 'junk percussion' collection which is used by MotET (MotET 2015), one of the bands which make up part of the Artifactory musical community.

Figure 6. Junk percussion rig.

Music Technology has been a strong pillar of activity of the Artifactory since its inception; it was a key interest for several of the founding members; and remains well represented in the governing committee and membership at large to this day. This bringing together of electronic music and maker culture at the Perth Artifactory happens at many levels, from amateur to professional. For example, the current Chairman is also the Chief Technology Officer of music software production company FXpansion (FXpansion); other members have been part of community reach groups with a music focus; instrument makers and so on, with chance and deliberate cross-fertilization between all. It has shown that there are many benefits to both makers and musicians in sharing resources and a diversely skilled community.

3. MUSIC GROUP

Current members of the Perth Artifactory involved in the music group include SKoT McDonald, Sam Gilles, Tim Gilchrist, Andrew Fitch, Nathan Thompson, Meg Travers, Rhys Channing, Josten Myburgh, Steve Paraskos, Stuart McDonald and Jean-Michel Maujean.

4. OTHER MEMBERS

Apart from the music group, members of the Perth Artifactory who have contributed to the projects listed above include:
Wayne Osborne (Lasers for NoizeMaschin!!)
Brett Downing (Construction of Arcophone)
Doug Bateman (Arduino U)
Brendan Ragan (Stage sound system)

5. MORE INFORMATION

http://artifactory.org.au

6. REFERENCES

The Perth Artifactory. *NoizeMaschin!!*. Retrieved 28th September 2015, available online at URL noizemaschin.artifactory.org.au/

nonlinearcircuits. *Analogue Modular Synthesizers*. Retrieved 28th September 2015, available online at URL nonlinearcircuits.com/

Perth International Arts Festival. *Hack the Festival*. Retrieved 28th September 2015, available online at URL 2015.perthfestival.com.au/Whats-on-by-Genre/Theatre-and-Circus/Hack-the-Festival

Make Magazine. *Arcophone Mark II Debuts*. Retrieved 28th September 2015, available online at URL makezine.com/2011/03/14/arcophone_mkii_debuts/

Tim Barrass. *Mozzi*. Retrieved 28th September 2015, available online at URL sensorium.github.io/Mozzi/

MotET. *So come down to the lab…*. Retrieved 28th September 2015, available online at URL www.motet.net.au

FXpansion. *FXpansion*. Retrieved 28th September 2015, available online at URL www.fxpansion.com/