Visualising the Sonic Environment

Lindsay Vickery
School of Music
Edith Cowan University

ABSTRACT

This paper describes score creating software implemented in the visual programming language MaxMSP, the *Lyrebird: Environment Player*. *Lyrebird* analyses and visualises significant features of a sonic environment as a graphic score, that is scrolled from right to left across the computer screen. The aim of the software is to create scores from field recordings with enough semantic detail to allow performer(s) to interact with the sounds of the environment. The form of interaction is open to the performer, but is envisaged to include precise emulation of the auditory features of the recording, improvisation informed by them or a combination of both. The software was specifically designed to partially remove the composer from the loop by not requiring any further specification or intervention in any performance than the generation of the score itself. It was also designed as a near-realtime tool, in which the environment of place and time of the performance could be reflected contemporaneously.

It is part of a project to construct processes allowing for data interchange between visual and sonic media: to create a continuum in which sound could be visualized and then resonified by both live performers and digital means (Vickery 2014b, 2015).

1. CONTEXT

Emulation of the sounds of the natural environment may be one of the earliest manifestations of music. Weiss argues that the representational/abstract nature of sound is one of "central ontological" debates in sound arts (Weiss 2008:11). Examples of the evocation of nature can be found throughout the history of Western Classical Music from 16th Century vocal music (O'Callaghan 2015) to the works of Messiaen (Harley 2008). The representation of natural environments in art music was uniquely altered, however, by the introduction of recording technology and digital analysis allowing sounds to be examined in increasing detail.

Slavoj Žižek claims, "technology and ideology are inextricably intertwined" (2000: 39) and in the emergence of works for acoustic instruments with field recordings this is certainly the case. The earliest non-anthropogenic sound was recorded as early as 1889 (Ranft 1997), but it was not until Pierre Schaffer's establishment of *Musique Concrète* in 1950 (Palombini 1993, Stock 2014) that "field recordings" became the materials of Art Music. Musique Concrète lifted the constraints of the metrical/chromatic grid imposed by traditional music notation, but the manipulation of recordings remained shaped by the intentions of the composer.

The erasure of the "grid" contributed to a re-evaluation of acoustic instrumental performance, in the "Musique Concrète Instrumentale" of Lachenmann that sought to illuminate "instrumental sounds as mechanical processes" (Lachenmann 1996: 212). Together with the other developments such as Cage's *Indeterminacy* (Cage 1961) and the emergence of *Free Improvisation* (Bailey 1993) a new space was opened for the emulation of natural sounds with "extended techniques" that exponentially expanded the timbral pallet of acoustic instruments. Alvin Lucier's (1931) *(Hartford) Memory Space* (1970) is a seminal work in this respect.

> *Go to outside environments (urban, rural, hostile, benign) and record by any means (memory, written notations, tape recordings) the sound situations of those environments. Returning to an inside performance space at any later time, re-create, solely by means of your voices and instruments and with the aid of your memory devices (without additions, deletions, improvisation, interpretation) those outside sound situations.*
>
> *(Lucier and Simon 1980).*

By the 1970s, R. Murray Schafer (1933), Hildegard Westerkamp (1946) and Barry Truax (1947) created the Vancouver World Soundscape Project (1972) at Simon Fraser University, identifying field recordings as art-works in themselves. This was perhaps as a response to the resurgence of the Environmental Movement following the publication of modern foundational texts such as Rachael Carson's *Silent Spring* (1962) (dealing with the effects of man-made pollutants on wildlife) and Paul Ehrlich's *The Population Bomb* (1968) – (concerned with the impact of the exponential growth of the human population). The term "Soundscape" was invented in 1967 by Schafer (Schafer 2006) and his colleague Westerkamp noted that compositions based on soundscape recording should be "rooted in themes of the sound environment" (Westerkamp, 2002: 53). This is an important distinction, elevating the structure and morphology of natural sounds beyond the manipulations of human-derived aesthetics and formed the basis for what was later termed Eco-structuralism which states that in such works "structures must be derived from natural sound sources" and "structural data must remain in series"(Opie and Brown 2006).

Lucier's work *Carbon Copies* (1989) in which performers imitate the sounds of any indoor or outdoor environment (albeit pre-recorded), "as exactly as possible, without embellishment" in a range of configurations: recording alone, instruments emulating the recording, instruments emulating the recording with the audio muted for the audience, instruments emulating the recording from "memory or by freely imagining the sounds continuing" (ibid) was a direct influence on *Lyrebird*.

Although Westerkamp's *Fantasie for Horns II* for french horn and tape (1979) predates *Carbon Copies*, it is a vehicle for an instrumentalist performing "common practice" musical gestures: chromatic temperament, harmony and metrical rhythm. Crucially there is no attempt to emulate the unmetered and freely tempered structures in the accompanying recording.

The shift from analog to digital recording also provided crucial new tools for analyzing field recordings with a grid finer than that of human perception. The spectrogram allowed for the visual representation of sonic events that were extremely difficult to capture with traditional notation, and for a much more precise emulation of continuous timbral features through "instrumental synthesis" (Grisey 2000).

The *Lyrebird Environment Player* sits at the nexus of these technologies/ ideologies:

- continuous unedited field recordings;
- visual representation of the field recordings is created through spectral analysis;
- the performer is expected to emulate the field recording as closely as possible using appropriate extended techniques;
- improvisation may be used to create sounds that fall within the context of the field recording.

2. TECHNIQUE

The conceptual basis for *Lyrebird: Environment Player* is straightforward: the performer is provided with a scrolling visualisation of an environmental recording 12 seconds in advance of it sounding, so that the recording is sounded synchronously with its visualised auditory features as they reach the left side of the screen. This arrangement creates continuous spatial representation of imminent events in

The work augmented techniques developed for the work *EVP* (2011) in which five performers were instructed to emulate the sounds in a pre-recorded audio collage using extended techniques with the aid of a scrolling score that showed relative pitch, duration and dynamics of the EVP samples in real-time (See Fig. 1).

A delay of 12 seconds, resulting in a scroll-rate of approximately 1.3cm/s, was chosen as a trade off between the degree of detail in the visualisation and legibility for the performer. A slower scroll-rate effectively "zooms out" the visualisation resulting in a lower resolution of sonic detail, whilst scrolling information on screen becomes increasingly hard to read as reaches the "fixation threshold of the human eye" (Picking 1997). Sightreading studies by Gilman and Underwood (2003) imply a maximal threshold rate for scrolling of about 3cm/s.

Crucially, this rate allows the performer to apprehend morphological detail of "human-scale" auditory phenomena – that is, within the limits of the mental chronometry of the human auditory system and response time.

The representation of a recording in *Lyrebird* is a compromise between a traditional musical score and a spectrogram. A spectrogram represents the energy of each frequency of a sound spatially, however it is poor at "demonstrating coindexation and segmentation due to the difficulty in illustrating differences in timbre" (Adkins 2008). This problem means that the spectrogram does not necessarily "look like" what we hear because "the sensory components that arise from distinct environmental events [are] segregated into separate perceptual representations" (Bregman 1990:44) by the human auditory system.

The musical score, on the other hand, represents actions to be undertaken in order to create a sound and yet this tablature provides enough information for musicians to "read" the score and mentally form a meaningful representation of its sound. The issues of dynamics, timbre and segmentation are, to a degree, accounted for in the score as a "bound-in" consequence of the actions represented in the score.

Lyrebird attempts to mimic the timbral separation of environmental auditory events by generating and

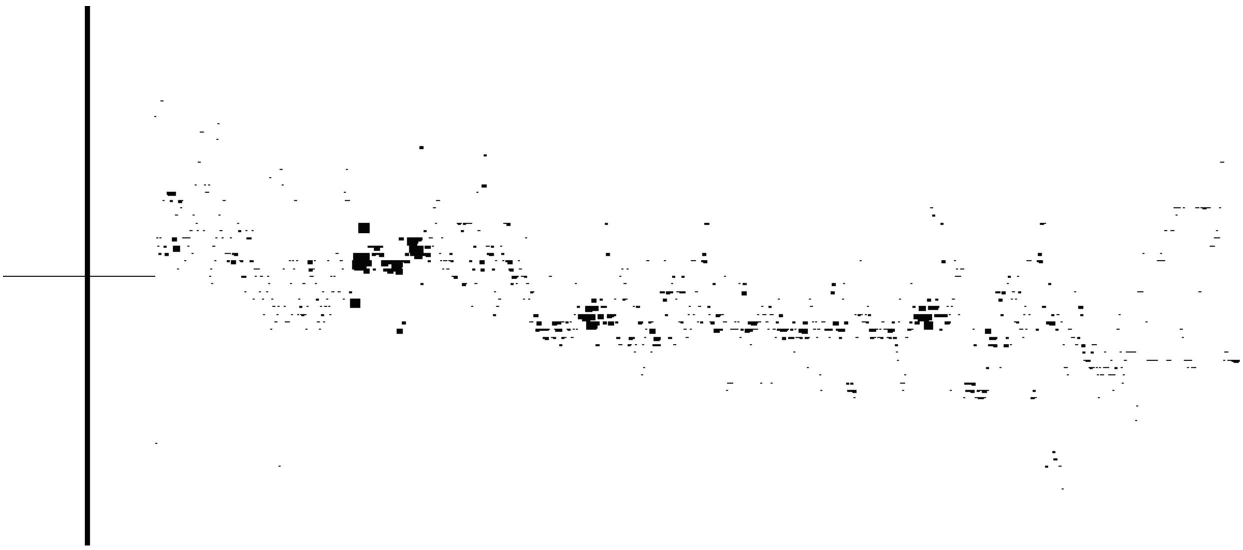

Figure 1. The scrolling scoreplayer for Lindsay Vickery's EVP [2012] showing visualized pitch and amplitude data.

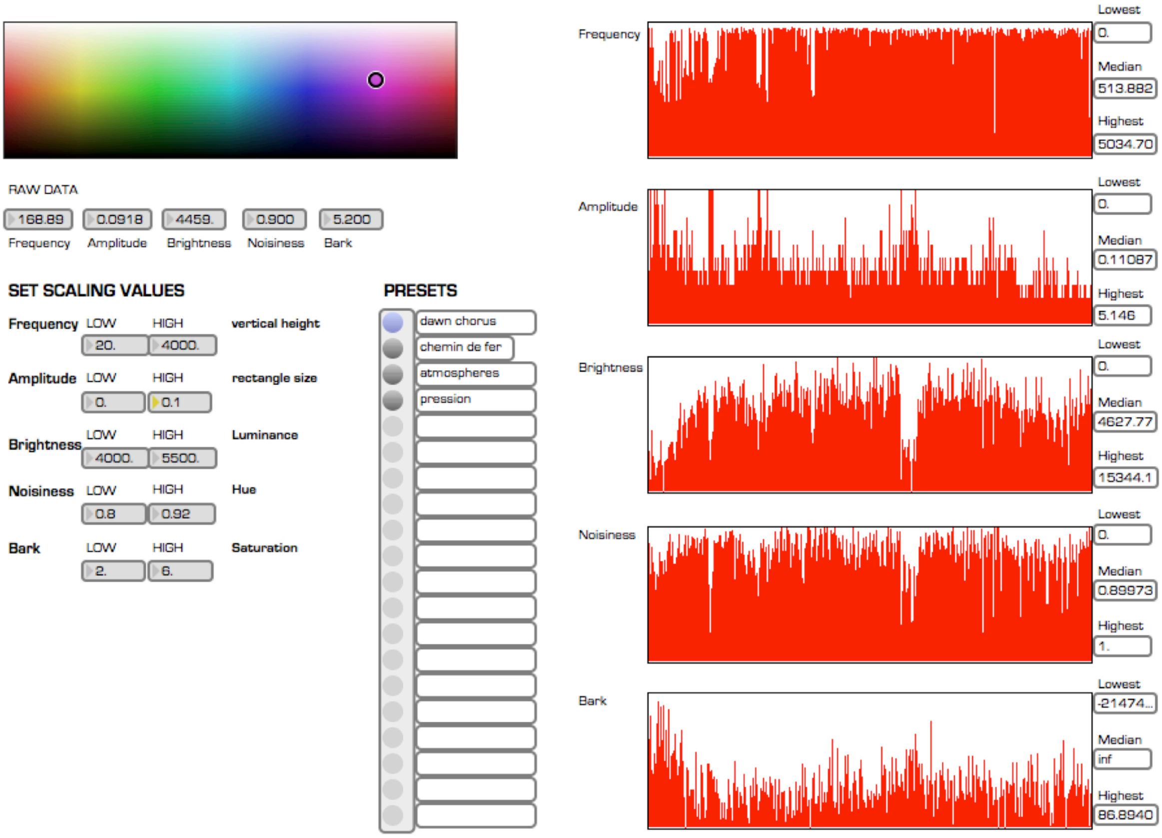

Figure 2. The Lyrebird: environment player analysis panel showing represented both raw value and scrolling multislider representation of data and scaling and preset options.

alternate form of spectrogram in which the strongest sinusoidal peak is represented spatially and coloured according to a simultaneous analysis of timbral features (brightness, noisiness and bark scale) of the whole recording at that moment. This process aims to assist the visual identification of auditory features.

The single strongest detected sinusoidal peaks detected by Tristan Jehan's `analyzer~` object are rendered as

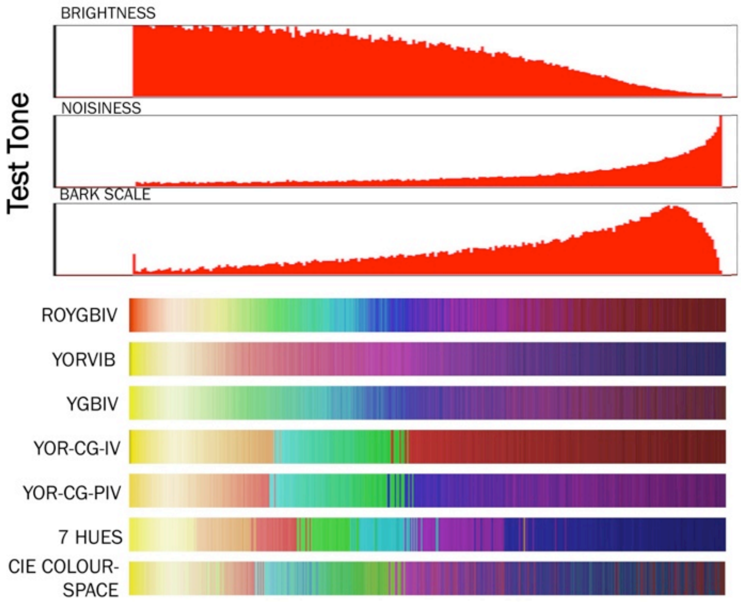

Figure 3. The spectra above depict a test tone of increasing brightness, noisiness and bark scale depicted by a variety of mappings.

rectangles drawn on a scrolling LCD object in MaxMSP. Each rectangle represents frequency as the vertical position, amplitude as size, and brightness, noisiness and bark scale data are mapped to hue, saturation and luminance values of the rectangle.

In the current version of this work, the median of 16 bark scale values (representing the deviations from expected auditory critical bands) is used. This presupposes that the median value refers to the same critical band as the strongest sinusoidal component. In future it may be possible to model this parameter more accurately.

Lyrebird incorporates an analysis panel (Figure 2.) that provides controls for the performer to view and scale data from the field recording. This allows for the performer to derive the most appropriate data for each individual recording and to "zoom" the visualization in or out on particular regions of activity in the recording.

To facilitate these decisions the data is represented both as a raw value and on a scrolling multislider displaying the its final scaled value, so that the performer may confirm that the scaling is capturing the full data range. The resulting colour is also displayed on a colour swatch. In the analysis panel, the performer may store the scaling values of up to 20 recordings.

Lyrebird allows for a range of mappings of timbral brightness to hue. The spectra in Fig. 3 depict a test tone of increasing brightness,

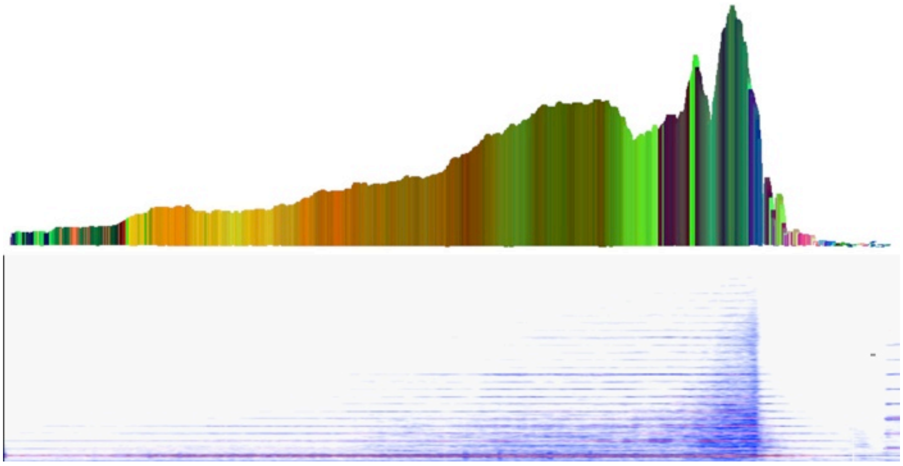

Figure 4. One of the crescendo F#s from the clarinet part of Messiaen's *Abîme des Oiseaux* represented as a spectrogram and the Lyrebird Environment Player.

noisiness and bark scale depicted by a variety of mappings.

Figure 4. illustrates a simple example in which one of the long-crescendo F#s from the clarinet part of Messiaen's *Abîme des Oiseaux* is shown represented as a spectrogram (using Chris Cannam's *Sonic Visualiser* software) (Cannam, Landone & Sandler 2010)) and the *Lyrebird Environment Player.* This example illustrates the representation of continuous timbral and amplitude changes over the duration of the note.

The example demonstrates the advantages of binding amplitude and timbre into a single form of representation in order to convey information about the sonic event, rather than representing every component frequency of the sound. This form of representation also addresses the problem that continuously evolving parameters such as timbre and amplitude, and the depiction of complex indeterminate environments such as those found in nature, are poorly captured by traditional Western music notation.

3. REPRESENTATION

The musical score is a time critical form of visualisation in which there is a strong imperative to employ symbols that signify sonic events with maximal efficiency and semantic soundness. *Lyrebird* draws on the concept that "pitch metaphors, while culturally diverse, may be based upon basic underlying mappings, stemming from bodily-based inter-modal interactions with the physical environment" (Eitan & Timmers 2010:407).

The simplest and perhaps least contested of these mappings is the vertical spatial depiction of frequency in which higher frequencies are also vertically higher on the page. The latent mapping of frequency to spatial height is demonstrated pan-culturally (Eitan and Timmers 2010:419) and in infants as young as 1 year old (Wagner, Winner, Cicchetti, and Gardner 1981).

Walker has proposed that such cross-modal correspondences are ordered in clusters. Walker claims "the same core correspondences should emerge whichever sensory feature is used to probe them, confirming that the en bloc alignment of the dimensions is context invariant" (Walker 2012:1806).

The concept of cross-modal correspondence also informs the mapping of louder amplitude to larger rectangle size in *Lyrebird.*

Application of cross-modal principals to colour is more problematic because of the difficulty of establishing a meaningful mapping of bright and dark colours. Whereas sound is mapped in a broadly linear fashion with the cochlea capturing frequencies continuously from high to low, the eye 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. 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 (Fig. 5).

CIELAB colour space (Figure 6.) attempts to mimic the nonlinear response of the eye by modelling cone responses. The notional colour spectrum provides a

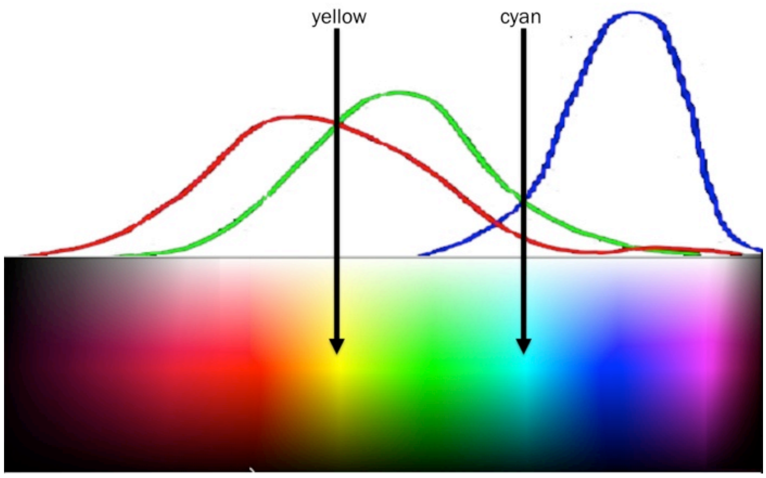

Figure 5. Light frequencies between the colour-sensitivity responses of the - red, green and blue cone cells (such as yellow, cyan and magenta) may appear brighter because they are perceived in regions where the sensitivity is attenuating.

palette from which colours representing sonic features or instruments might be chosen in a musical score. For most people this chart appears segmented into families of similar hue (yellow, orange, tan, green-blue etc) and distinct but related hues may lend themselves to the

Figure 6. A notional colour spectrum based on human visual perception from white to black (based on CIELAB colour space (Hoffman 2003) and Bruce MacEvoy's Artist's Value Wheel (MacEvoy 2005).

representation of timbral variation within a sonic feature or instrument.

Lyrebird allows a range of timbre to colour mappings: ROYGBIV, YORVIB (bypassing the green-blue channel), YGBIV (bypassing the Red-Orange channel), YOR-CG-IV and YOR-CG-PIV (comprising three light-to-dark spaces), 7 Hues (YORGCMB) and CIELAB COLOUR-SPACE. The performer may experiment to discover which mapping is most suited to a specific field recording.

As sounds in nature are often arranged in frequency/timbre bands – frogs, large birds, small birds, crickets etc – some mappings prove to be more effective at representing a particular environment. The score may also be viewed against a black of white background (Fig. 7). *Lyrebird* does not at present support fully customised mappings.

Future versions of *Lyrebird* may incorporate recent research at *The Visual Perception and Aesthetics Lab* at the University of California Berkeley, suggesting that there is a high degree of correlation between mappings of colour-to-sound in non-synaesthetes. Griscom and Palmer have proposed that there are systematic relationships between colour and a range of musical phenomena including timbre, pitch, tempo, intervals, triads and musical genres (Griscom and Palmer 2012, 2013).

Colour provides a great potential for the formation of Perceptual Discriminability in a musical score. One obvious approach, for example, might be to employ a colour scheme that maximizes the distinctness of separate musical phenomena such as instruments, voices or sound sources. Similar requirements have been studied for the creation of data visualisation (Tufte 1990), transport maps (Green-Armytage 2010), and websites (Stanicek 2009). Recent research, however, has indicated strong perceptual correspondences between colour and a range of sonic phenomena (Prado-Leon, Schloss & Palmer 2011), suggesting there may be more intrinsic semantic value to be gained from colouring the score.

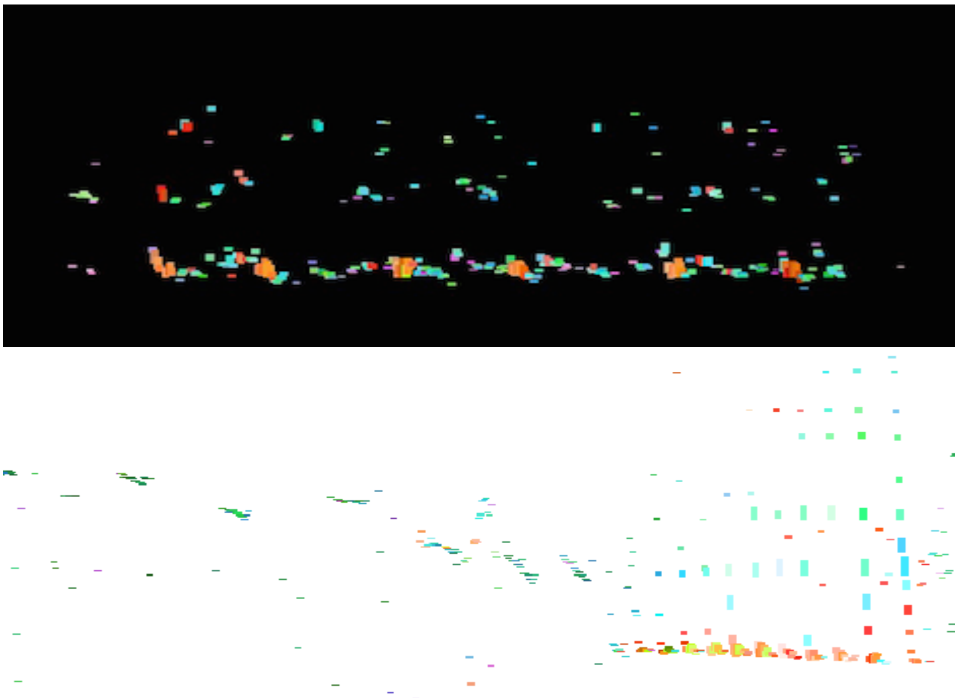

Figure 7. Screen shot of black background *Lyrebird* score from *Bullfrogs and Rainstorm* (excerpt above) and Screen Shot of white background *Lyrebird* score from *Whistlers and Crickets* (excerpt below).

4. DISCUSSION

In this project, the objective of interaction between a live performer and environmental sounds in works such as Lucier's *Carbon Copies* was broadened through the addition of a visual representation of the field recording. While not strictly "real-time" (recorded sounds are delayed by 12 seconds), this environmental sound scoreplayer allows performer(s) to engage with natural sonic environments in a site-specific manner, using field recordings and sonorous objects from the vicinity of a performance. It

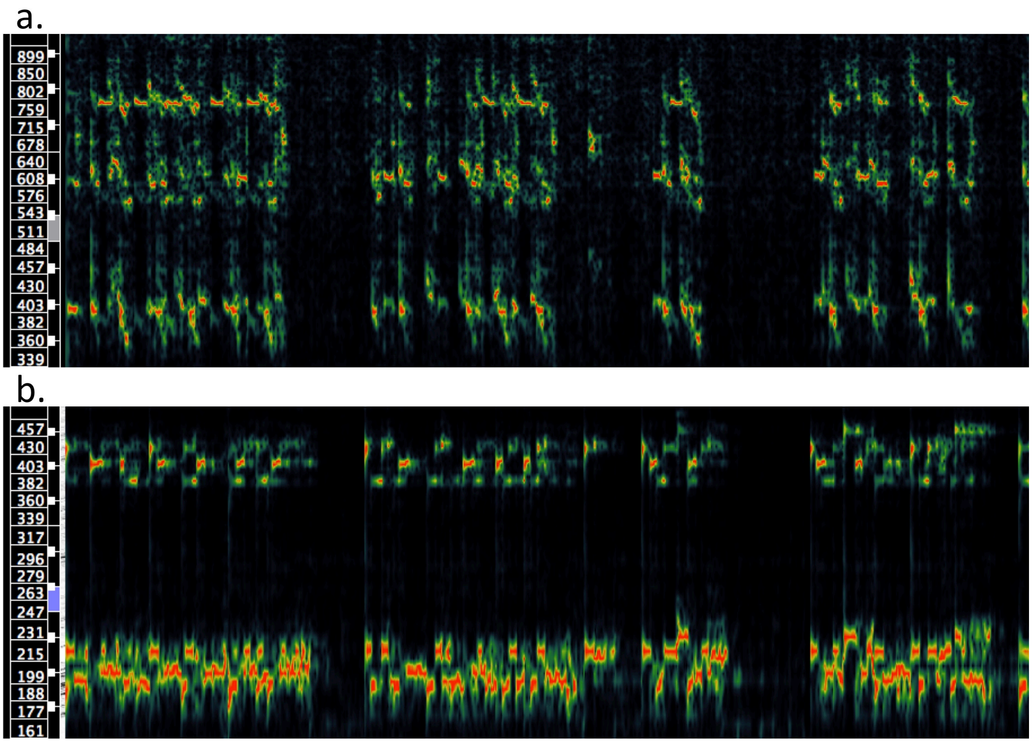

Figure 8: *Bullfrogs and Rainstorm:* comparison between (a) field recording 339-899Hz and (b) piano performance 161-457Hz.

provides a mechanism for the performance of, or improvisation around, significant sonic features from the natural environment.

The figures below explore the effectiveness of these objectives through the comparison spectrograms of the source field recordings and the performances of a number of musicians. The spectrograms depict comparable frequency ranges, although the performer spectrograms are uniformly an octave lower than those of the field recording. The titles of the recordings are those given by the original recordist Philip Kenworthy.

Figure 9 compares the field recording *Bullfrogs and Rainstorm* with a pianist's performance. The performer's relationship to *Lyrebird* is the most "score-like", in that the pitch, rhythmic and dynamic contours of the bullfrog croaks from the field recording are adhered to with a great deal of precision.

The task is perhaps simplified because the pitch range of the croaks is limited to about 3 semitones, however the spectrogram indicates that this method of synchronisation of the recording and the performance is effective in this instance.

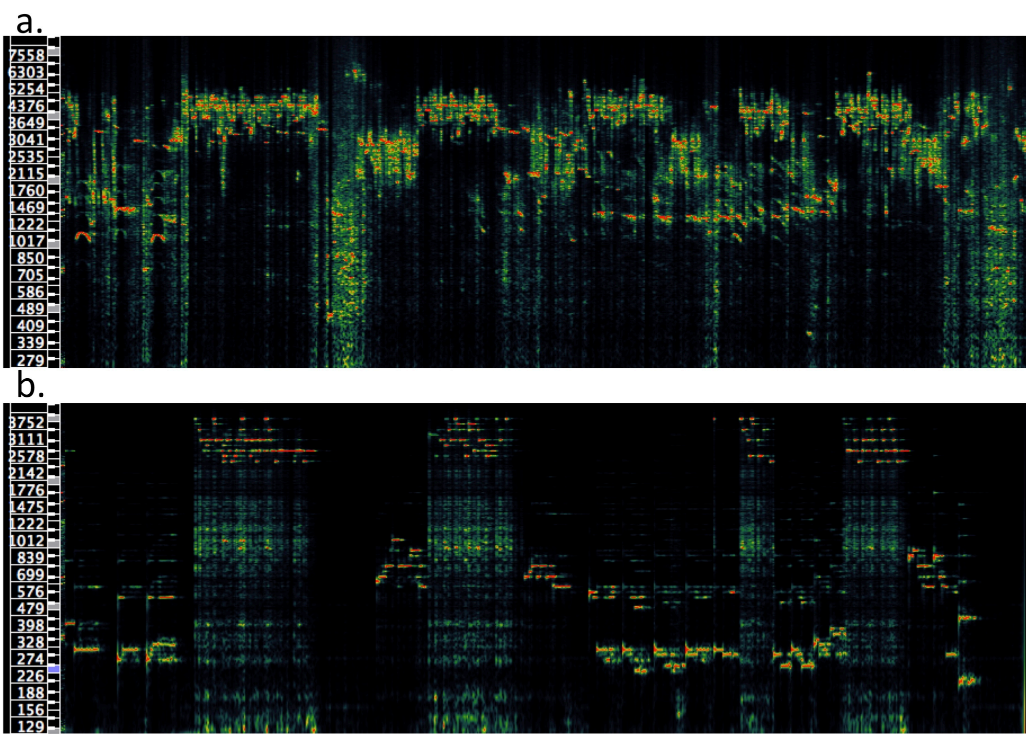

Figure 9: *Kookaburra Sunrise:* comparison between (a) field recording 279-7558Hz and (b) piano performance 129-3752Hz.

In a case where a more complex sound environment comprising a range of birdsongs is used, such as Fig. 10, the task is considerably more complex. Birdsongs are often too high pitched and rapid to be emulated entirely accurately. *Lyrebird* aims to provide information at a human scale showing the contour of extremely rapid birdcalls rather than precise detail. Here the same pianist aims to portray three bands

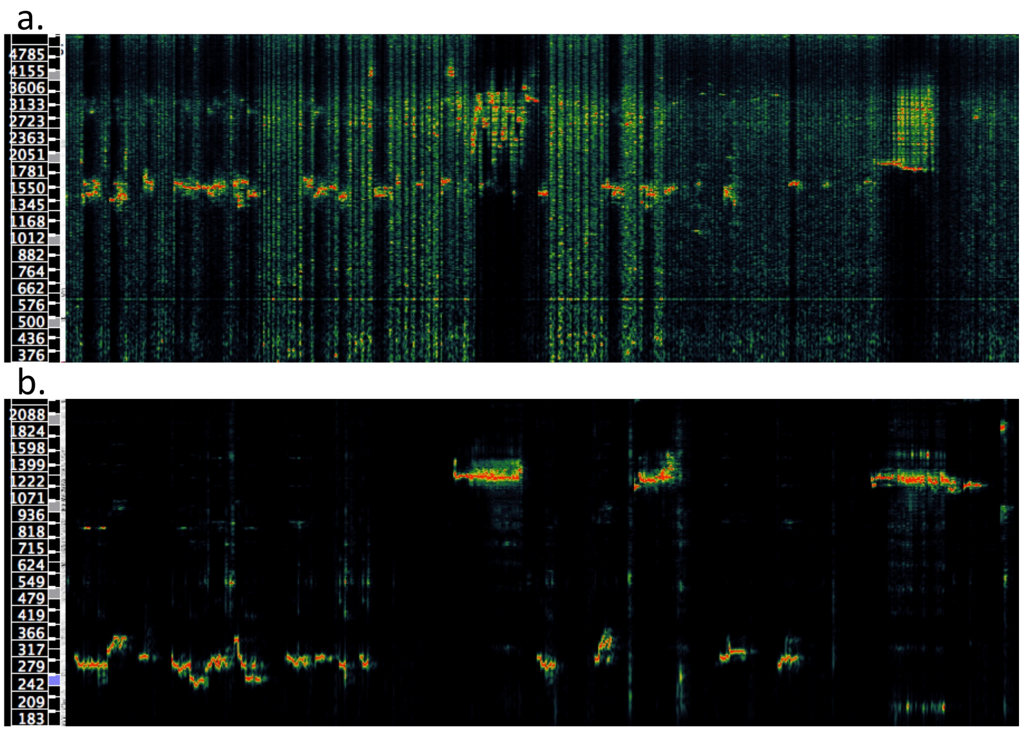

Figure 10: *Whistlers and Crickets:* comparison between (a) field recording 376-4785Hz and (b) bass clarinet performance 183-2088Hz.

of activity in different ranges and is successful until the bands have simultaneous activity (about half-way through the figure), at which point activity in the highest band is ignored.

In a complex environment such as the recording *Whistlers and Crickets* (Fig. 11) there is too much data for the performer to emulate the complete environment. Again the performer, this time on bass clarinet chooses specific features to emulate. A potential solution to the emulation of complex environments is the use of networked players displaying activity in different bands for a number of performers.

Lyrebird does not represent the sonic events on a grid to indicate frequency, and it is evident that the performer here "overshoots" the high frequency events, performing them about an octave too high in relation to the lower pitched layers.

In Fig. 12, the performer's aim was to interact with the recording through both emulation and improvisation. The field recording spectrogram shows the gradual fading of repetitive cricket sounds, followed by a prominent birdcall.

The percussionist emulates only the birdcall and then mimics the cricket sounds once they have ceased. The passage suggests that *Lyrebird* can provide and effective representation of the sonic environment that allows the performer to interact by taking over features from the recording.

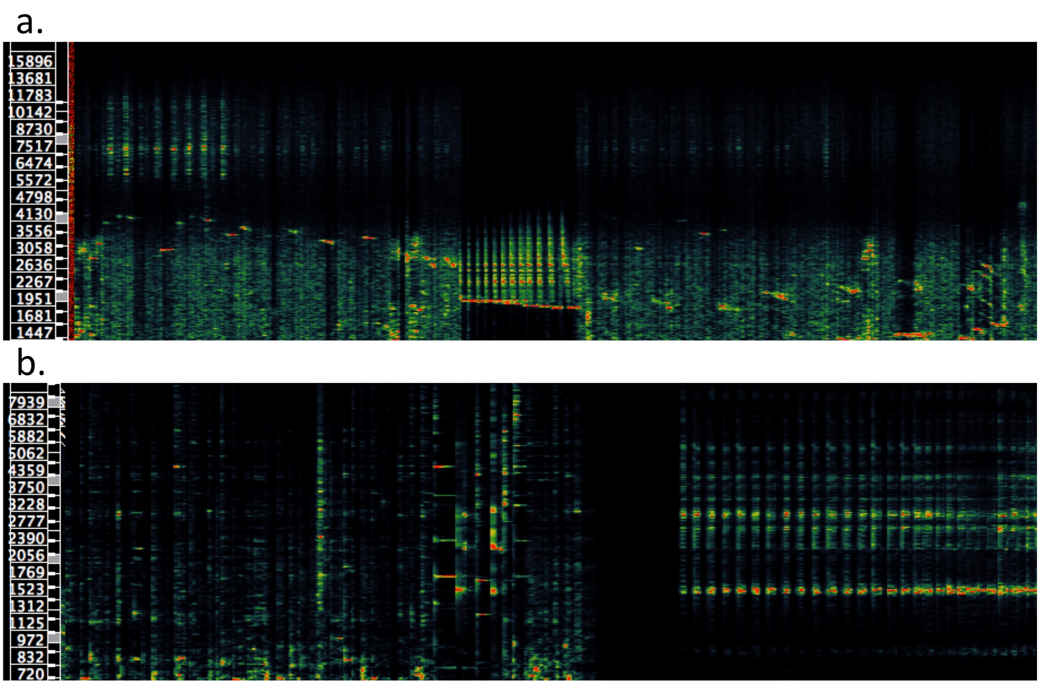

Figure 11: Whistlers and Crickets: comparison between (a) field 1447-15896Hz recording and (b) percussion performance 720-7939Hz.

Lyrebird allows for interaction with pre-recorded non-anthropogenic sound environments. The degree to which an interaction is meaningful is self-evidently dependent on the performer(s) abilities. Examples of this work can be heard at (https://lindsayvickery.bandcamp.com/album/works-for-instruments-and-field-recordings). However, unlike many musical experiences the potential for precise synchronisation of a performer with seemingly indeterminate sonic events arguably has an intrinsically interesting quality. The evaluation of the accuracy of performer emulation is something of an end in itself, and this includes both the degree of acoustic reproduction of the sounds (as demonstrated in Figures 9-11) and the performer's ability to "enter into" the soundworld of the recording through improvisation.

5. CONCULSION

Efficient and semantically sound notation is crucial for the development of effective notation for screenscores. The compromise between a traditional musical score and a spectrogram adopted in *Lyrebird: Environment Player* is part of an on-going project to explore means to better capture nuances of sound such as timbre, temperament and envelope morphology using shape and colour parameters (hue, saturation and luminosity).

The arrangement of displaying data before it is heard proves to be an effective means of preparing the performer(s) with spatial and timbral information comparable, in general terms, to that of traditional music score.

Future developments of the software may be further enhanced to reflect more complex or customised colour mapping. Bark scale values, that are currently averaged may be modelled more accurately across the entire frequency spectrum, for example. Other spectral parameters such as diffuseness may be exploited as potential bearers of timbral information. Higher order Auditory Scene and morphological features of spectra may be taken into account such as flux, contour and stream segregation, pursuing the goals of composers of Spectral music (Fineberg 1999). Such developments may allow for data to be assigned to particular instruments within an ensemble dynamically through interpretations based upon range, timbre within and polyphony.

Multiple scoreplayers have already been networked together, allowing ensembles of performers to interact with visualisations that focus of varied frequency, amplitude and timbral parameters of the same recording.

Lyrebird can also operate as a tool in the creation of fixed works using a single field recording, allowing for greater annotation of the score or to generate timbral information for an otherwise fixed score as is the case in *the miracle of the rose* (2016) (Vickery 2015).

A current project is seeking to develop means for communicating data to the more portable Decibel Scoreplayer (Hope et al 2015).

The examples of performances with *Lyrebird* suggest it can effectively convey useful information to the performer and act as a vehicle for live performers to engage with field recordings.

6. ACKNOWLEDGEMENTS

Lyrebird was commissioned by percussionist Vanessa Tomlinson for her Australian solo percussion program *Eight Hits*. The author and Tomlinson developed the work during her residency at the *Orpheus Instituut for Advanced Studies & Research In Music* in December 2013.

The field recordings discussed in section 5 were recorded by Philip Kenworthy.

Recordings of performances of *Lyrebird* were made as part of an Early Career Researcher Grant from Edith Cowan University.

7. REFERENCES

Adkins, M. 2008. The application of memetic analysis to electroacoustic music. http://www.mathewadkins.co.uk/article4/.

Bailey, D. 1993. *Improvisation: Its Nature and Practice in Music,* Da Capo p.83.

Cage, J. 1961."Indeterminacy". *Silence.* Middletown, Conn.: Wesleyan University Press.

Cannam, C. et al., 2010. "Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files", *Proc. ACM Multimedia 2010 Int. Conf. Firenze,* pp. 1467-1468.

Carson, R. 1962. *Silent Spring*. Houghton Mifflin.

Ehrlich, P. R. 1968. *The Population Bomb*. Sierra Club/Ballantine Books.

Fineberg, J. 1999. *Sculpting Sound An introduction to the Spectral Movement - its ideas, techniques and music*. Doctor of Musical Arts. School of the Arts Columbia University.

Gilman, E. and Underwood, G. 2003. "Restricting the field of view to investigate the perceptual spans of pianists". *Visual Cognition* 10(2) 201-232.

Green-Armytage, P. 2010. "A Colour Alphabet And The Limits Of Colour Coding". *Colour: Design & Creativity,* pp. 510 1–23.

Griscom, W. S., And Palmer, S. E. 2012. "The Color Of Musical Sounds In Non-Synesthetes". *12th Annual Meeting Of The Vision Science Society, Naples, Fl.*

Griscom, W. S., And Palmer, S. E. 2013. "Cross-Modal Sound-To-Sight Associations With Musical Timbre In Non-Synesthetes". *13th Annual Meeting Of The Vision Science Society, Naples, Fl.*

Grisey, G. 2000. "Did You Say Spectral?" [1998] Trans. Joshua Fineberg. *Contemporary Music Review* 19(3) pp. 1–3.

Hope, C. et al. 2015. "The Decibel Scoreplayer: Enriched Scores For The iPad". *Proceedings of the International Computer Music Conference 2015, Denton, Texas,* 314-317.

Lucier, Alvin, and Douglas S. 1980. *Chambers.* Middletown, CT: Wesleyan UP.

O'Callaghan, J. 2012. "Mediated Mimesis: Transcription as Processing*". Proceedings of the Electroacoustic Music Studies Network Conference, Stockholm.*

O'Callaghan, J. 2015. "Mimetic Instrumental Resynthesis". *Organised Sound* 20(2), pp 231-240

Opie, T. and A. R. Brown 2006. "An Introduction to Eco-Structuralism". In Proceedings of the 2006 International Computer Music Conference, New Orleans, USA, pp. 9-12,

Palombini, C. 1993. "Machine Songs V: Pierre Schaeffer: From Research into Noises to Experimental Music". *Computer Music Journal* 17(3):14-19.

Peacock , K. 1988. "Instruments to Perform Color-Music: Two Centuries of Technological Experimentation". *Leonardo*, 21(4), pp. 397-406.

Prado-Leon, L. R. et al. "Color, Music, And Emotion In Mexican And US Populations". *New Directions In Colour Studies*. Amsterdam: John Benjamins.

Ranft, R. 1997. "The Wildlife Section of The British Library National Sound Archive (NSA)". *Bioacoustics: The*

International Journal of Animal Sound and its Recording 7(4) 315-319

Schafer, R. M. 2006. “I Have Never Seen a Sound”. *Environmental & Architectural Phenomenology* 17(2).

Stanicek, P. 2009. *Color Scheme Designer*. http:// http://colorschemedesigner.com.

Stock, J. 2014. *The art of the commonplace: Found sounds in compositional practice*. PhD. City University of New York,

Tufte, E. 1990. *Envisioning Information*. Graphics Press.

Vickery, L. R. 2014. “Exploring a Visual/Sonic Representational Continuum”.. *Proceedings of the International Computer Music Conference 2014*. 1: pp. 177-184.

Vickery, L. R. 2015. “Through The Eye of the Needle: Compositional Applications For Visual-Sonic Interplay”. *Proceedings of the Australasian Computer Music Conference 2015, Sydney*. pp. 126-133.

Weiss, A. 2008. *Varieties of Audio Mimesis: Musical Evocations of Landscape.* Errant Bodies Press, Berlin.

Westerkamp, H. 2002. “Linking soundscape composition and acoustic ecology”. *Organised Sound* 7(01).

Zizek, S. 2000. *The Art of the Ridiculous Sublime*. Seattle Washington: Walter Chapin Centre for the Humanities: Occasional Papers 1, University of Washington Press.