HYBRID REAL/MIMETIC SOUND WORKS

Lindsay Vickery
Edith Cowan University
l.vickery@ecu.edu.au

ABSTRACT

This paper describes a project to construct a process allowing for data interchange between visual and sonic media: to create a continuum in which sound could be visualized and then resonified through by both live performers and digital means.

A number of processes to aid this visualisation/sonification "ecosystem" were developed. Software was created to create scores based on sonic features of "field recordings" through spectral analysis by rendering the frequency of the strongest detected sinusoidal peak of a recording vertically and its timbral characteristics by luminance, hue and saturation on a scrolling score. Along similar principals a second process was developed to generate a realtime score using graphical symbols to represent detected accents in "found sound" speech recordings. In the other direction software was built to render greyscale images (including sonograms) as sound and a second iteration to generate audio from detected analysis parameters.

The imperfections in the various transcription processes are intriguing in themselves as they throw into relief the distinctions between the various forms of representation and in particular the timescales in which they are perceived. The implied circularity of processes also opened the potential for re-interrogation of materials through repeated transmutation. This discussion explores these implications in the context of the analysis of field recordings to generate visual representations that can be resonified using both performative (via notation) and machine (visual data-based) processes, to create hybrid real/mimetic sound works through the combination (and recombination) of the processes.

1. INTRODUCTION

This paper describes a project to construct means to interchange data between visual and sonic media: to create a continuum in which sound can be visualized and then resonified. *Nature Forms II* (2015) – and the work that has led to it - is discussed as a vehicle for exploring the possibility of recursive re-interrogation of a field recording through visualization and resonification/resynthesis via machine and performative means.

In the work's predecessor, *Nature Forms I*, a score comprising manipulated images of organic shapes derived from photographs of trees, plants and rocks, was simultaneously sonified by performers and software. Three performers and software "read" from the same scrolling score on networked laptops

Four contrasting forms of reading/sonifcation were involved: machine sonification in which spatial position and colour of features from the image were more or less precisely rendered; tablature in which spatial position and colour were recast against the geography of a specific instrument; semantic reading in which the performer's understanding of notational conventions informed the outcome; and aesthetic reading in which the performer's understanding of the conventions of sonic representation were drawn upon [1].

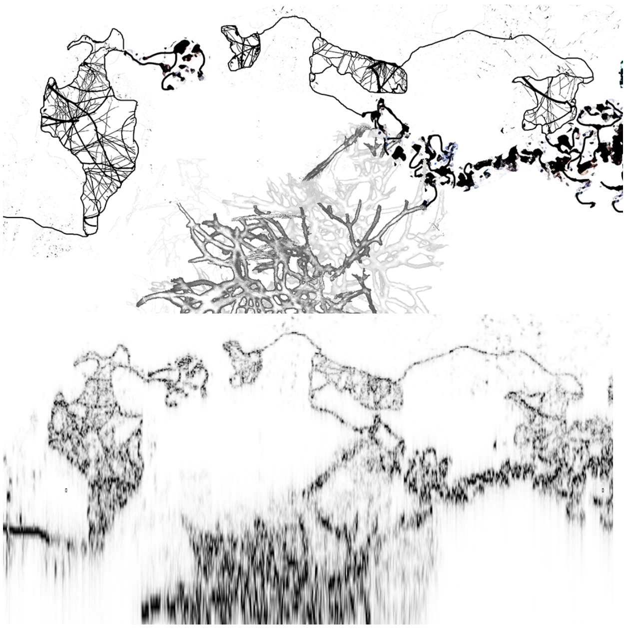

Figure 1. Comparison between an excerpt from the score of *Nature Forms I* [2014] (above) and a sonogram of its sonification (below).

In *Nature Forms II* a range of forms of representation are also explored including semantic graphical notation, percussion notation and a hybrid form of sonogram notation. The concept of multiple notations was also emphasised in another work, *Sacrificial Zones* (2014) in which the performer reads from a rhizomatic score - the notation moves along interconnected vertical and horizontal pathways - that crossfades between five layered images, each notated in a manner corresponding to a different form of visual representation of sound: non-semantic graphical

notation, semantic graphical notation, traditional notation, proportional notation and a spectrogram.
The score confronts the performer (and vicariously the audience for whom it is projected) with the variation in freedom and constraint presented by a range of forms of notational representation. The rhizomatic and layered procedure for rendering the score allows for multiple versions of this work, emphasising different aspects of the relationship between varied notations of the same musical object [2].

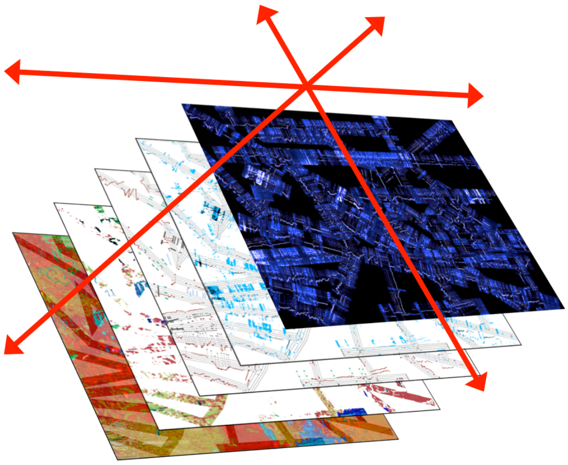

Figure 2. Layers of different visual representation of sound in *Sacrificial Zones*.

This conceptual basis assumes that visual forms of musical representation may be considered to occupy a continuum, between the spectrogram (a precise frequency/time/amplitude representation of sound), proportional notation, traditional notation, semantic graphical notation, non-semantic graphical notation [3]. It is also assumed that notation has inherent semantic implications as a consequence of a degree of ‘weak synaesthesia’ [4, 5] or cross-modal activation that is present in the population at large, and therefore that graphical symbols can elicit meaning through the inherent semantic qualities of their shape and colour, and that consideration of these qualities is crucial for the development of effective and efficient notation for screen scores [1, 2, 3].

2. VISUALISATION PROCESSES

A number of forms of notation/visualization are employed in *Nature Forms II*. This most literal is the process I have previously termed the “spectral trace” (Vickery 2014b) in which notation is drawn directly onto the spectrogram. This approach was used in the work *acid fury* (2015) in which colour-coded parts for the eight instruments were made from direct transcriptions of a spectrogram of the recordings (Fig. 3.).
Nature Forms II employs this process to represent features of the field recording to be performed by an ensemble of clarinet (orange), viola (red), and cello (green) (Figure 4.). The frequency/amplitude morphology of features of the field recording is communicated to the performer by extracting shapes directly from the spectrogram.

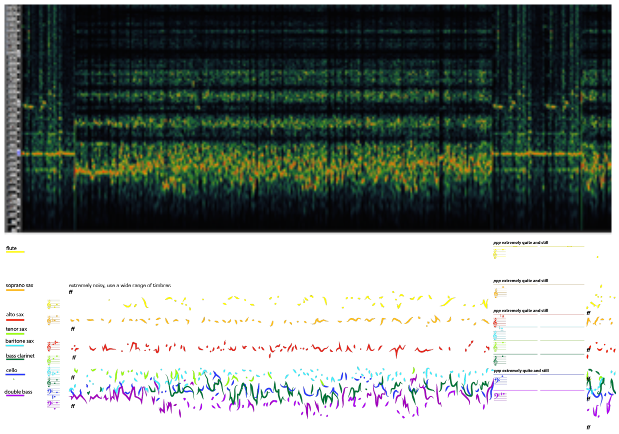

Figure 3. Spectrogram of found-sound recording (above) and annotated graphical transcription (below) of *acid fury* [2015] (excerpt).

A system presenting graphical symbols on a staff that is proportional both horizontally and vertically was created to solve the problem each pitch in a spectrogram occupies a distinct vertical spatial position. The system was developed for the work *here, apparently, there was time for everything* [2015], and attempted to more-or-less retain the topographical layout of the traditional stave, while adding coloured lines to indicate non-natural notes (Fig. 5.). This approach was eventually abandoned in favour of annotation of the score with pitch and articulation information in reference to each spectrally traced graphic.

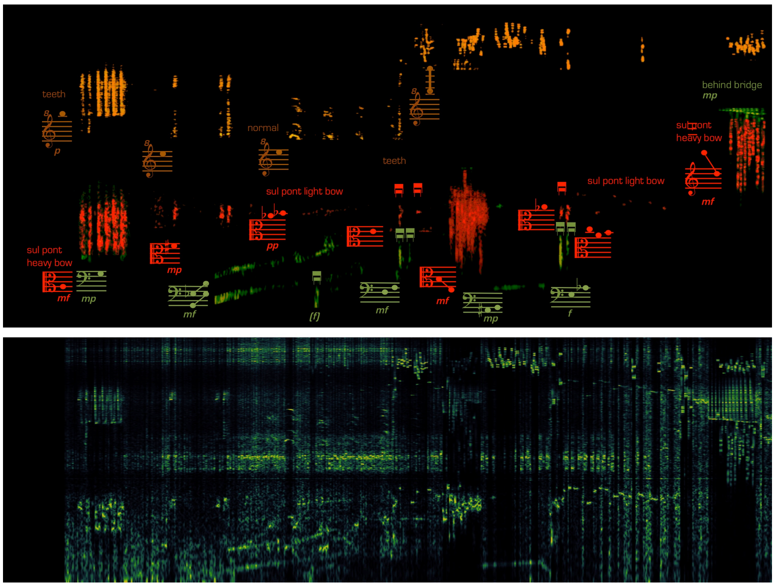

Figure 4. “Spectral Trace” notation from *Nature Forms II* (above), source spectrogram (below).

A second approach drews on the concept and techniques developed in *EVP* (2011) and *Lyrebird: Environment Player* (2014) in which the amplitude, frequency, brightness, noisiness and bark scale[1] of the single strongest

[1] The median of 16 bark scale values (representing the deviations from expected critical bands) is used. This presupposes that the median value refers to the same critical band as the

detected sinusoidal peak in a recording is represented by the vertical height, size, luminance, hue and saturation of rectangles drawn on a scrolling LCD object (in this case *jit.lcd*).

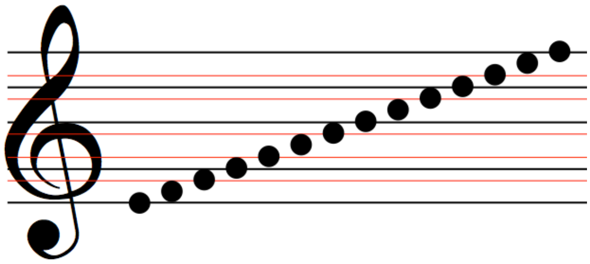

Figure 5. Vertically proportional stave showing a chromatic scale from E3 to F4.

Data is derived from Tristan Jehan's `analyzer~` [6] object in realtime allowing for the scoreplayer to visualise timbral features of the recorded sound.[2] The visualised score depicting the principal features of a source recording is scrolled from right to left across the computer screen and playback of the source recording is delayed (12 seconds in this work) to allow the performer to see a visualisation of the sounds before they appear (Figure 6.)

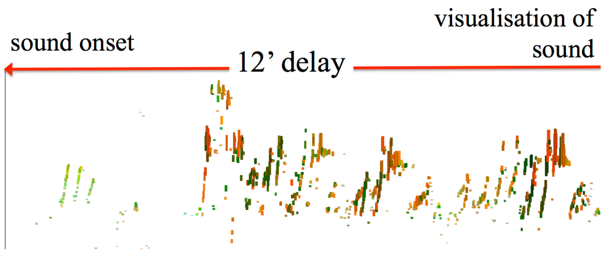

Figure 6. The scrolling scoreplayer for *Lyrebird: environment player* [2014] showing visualized pitch, amplitude and timbral data.

In the work *murmurs trapped beneath the bark* [2014], an this idea is elaborated through the use of a processed recording of a clarinet improvisation as the source audio. Its ambiguous resemblance to real-world natural sound led me to term this an "artificial field recording". In *murmurs* the score produced by *Lyrebird* is also visually processed in *Illustrator* to further the analogy with the idea of an artificial recording (Fig. 7.). It was this work, and this process that suggested the possibility of creating hybrid real/mimetic sound works interchangeably combining field recordings and their machine and human emulations.

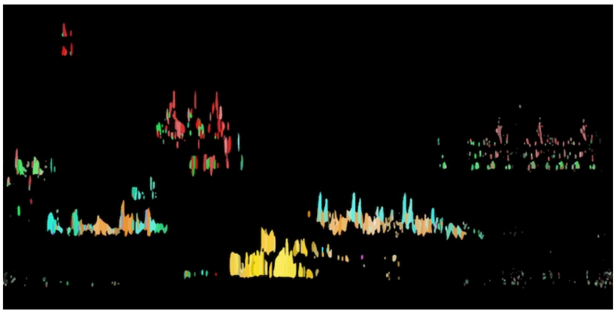

Figure 7. Detail from the score of *murmurs trapped beneath the bark* [2014].

The "lyrebird" generative score process was also employed in *Nature Forms II*, to create visualisations of the field recording. The *lyrebird* visualisations are scalable allowing for renderings of the score that focus upon high, middle and low frequency bands of the field recording.

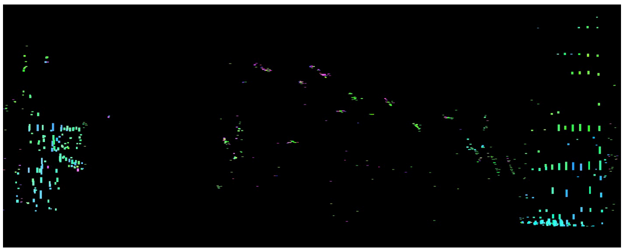

Figure 8. Detail from the high frequency "lyrebird" score of *Nature Forms II* [2015] (excerpt).

In the work *the miracle of the rose* [2015] these ideas were further elaborated combining 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]. A collage of time-stretched recordings of the text by Australian/French artist Emmanuelle Zagoria was used as the underlying structure of the work. The spoken phrases were transcribed for the two percussionists into gestures exploring their cadence and timbre via varied instruments and notational approaches.

In Fig. 9 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) for a bowed cymbal gesture that follows the envelope of the fixed media recording. The notation for this figure was again created using *Lyrebird: environment player* software.

In the upper half of Fig. 9 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 curved 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.

strongest sinusoidal component. In future it may be possible to model this parameter more accurately.

[2] A version using externals by Alexander Harker [7], using spectral centroid, spread and skewness is currently being trialed.

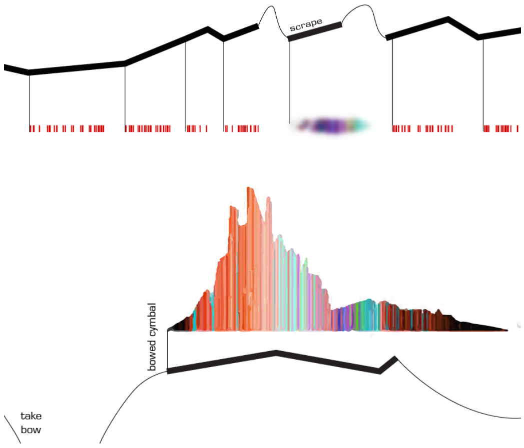

Figure 9. Excerpt from the score of *the miracle of the rose* [2015].

In Fig. 10. 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.

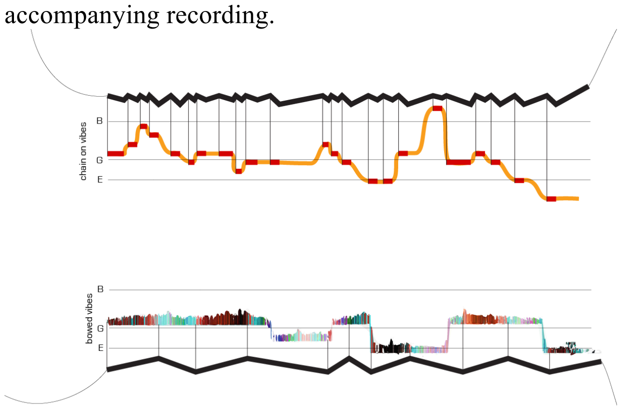

Figure 10. Excerpt from the score of *the miracle of the rose* [2015].

The final form of visualization in *Nature Forms II* is derived from the work *Semantics of Redaction* (2014) in which notation in generated (Figure 11.) by using accents detected in a speech recording in real-time to generate graphical symbols of varying vertical position, size and colour, determined by the frequency, amplitude and timbre of a speech recording at the accent point. In *Nature Forms II* the software detects accepts created by sharp attacks such as the chirping of crickets. This method was employed as the paradigm for the percussion notation for *Nature Forms II.*

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. [8]

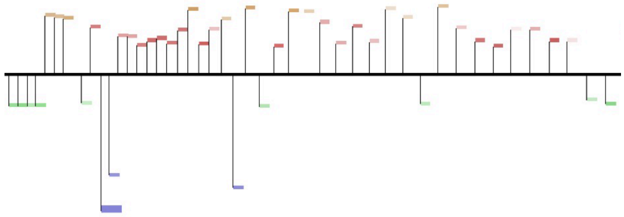

Figure 11. Percussion part for *Nature Forms II* created using the generative score software *Semantics of Redaction* (excerpt).

In particular the works allow for the contrast and interaction of instrumental and machine forms of sonification.

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. Figure 12. shows an excerpt of the workings for the score of *the miracle of the rose*, indicating how the score was developed in reference to spectrograms of the time-stretched recordings. The blue lines below the spectrograms indicate which of the recordings is sounding at any moment and therefore which form of notational representation is employed to represent the sound of the fixed media.

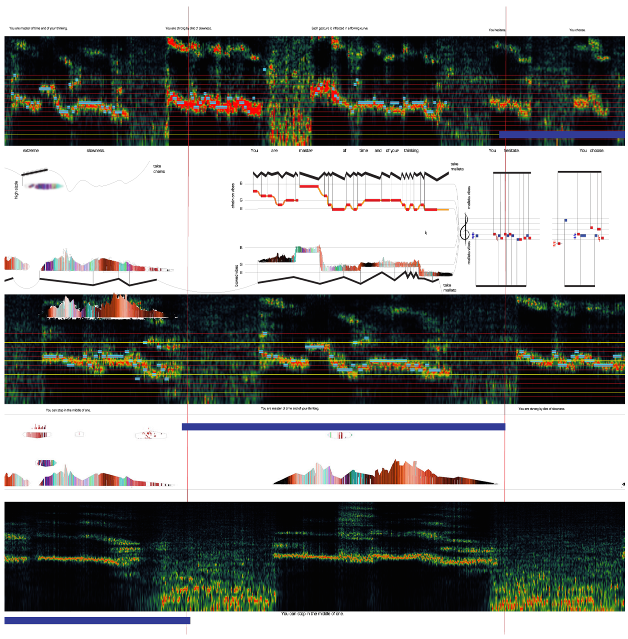

Figure 12. Score workings from *the miracle of the rose* [2015] showing the forms of notational used to represent the sound of the fixed media.

3. SONIFICATION PROCESSES

A number of approaches to sonfication/resynthesis are used in *Nature Forms II*. The first is an additive synthesis approach using a patch called *Sinereader* [1] developed in MaxMSP to resonify greyscale spectrogram images (Fig. 13.). 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.

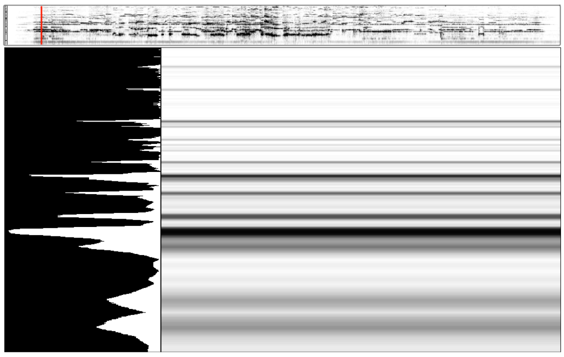

Figure 13. The *Sinereader* patch The complete spectrogram with a "scrollbar" indicating progress through the image is displayed at the top of the image, the greyscale value of each vertical pixel in a one pixel segment is displayed on the bottom right and the resulting amplitude is displayed on the bottom left.

Processes were also developed to resynthesise sounds using spectral analysis. In the first process, the strongest sinusoidal component detected each 40ms of the recording were resynthesized with a sinewave that was then ring-modulated according to the currently detected brightness of the recording (Figure 13.). The aim was to retain the amplitude of sonic features of field recording while maintaining and equivalent brightness.

Figure 13. Spectrogram of Ring Modulation synthesis from *Nature Forms II* (excerpt).

This method is also used in portions of the performance to ring modulate the instruments via live processing. In order to mimic the parametric brightness of the field recording the brightness of the live instruments is subtracted from that of the field recording to derive a ring modulation value.
Subtractive synthesis was also employed by using frequency and amplitude data detected in the recording to bandpass filter white-noise (Figure 14.). The result of these processes are sonic abstractions sharing morphological traits, but sonically distant from the source material

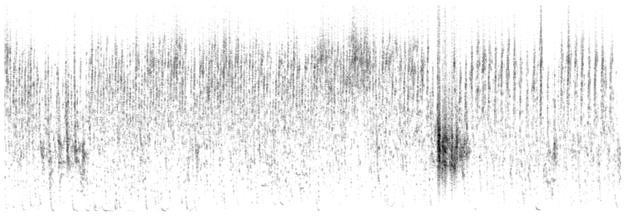

Figure 14. Spectrogram of subtractive resynthesis from *Nature Forms II* (excerpt).

At the opening of *Nature Forms II* "coloured noise" [9] performed by the instrumentalists is gradually shaped into the sonic structure of the field recording using subtractive synthesis, and then cross-faded with the source recording.
A similar process is used by Peter Ablinger in some of his "Phonorealist" works in the *Quadraten* series [8], in which spectral analysis data from recordings is "reconstituted in various media: instrumental ensembles, white noise, or computer-controlled player piano" [10]. A key issue at the heart of *Quadraten* is representation or analogy made between "real" sounds and their reconstituted counterparts. The transmutation through different forms of synthesis causes a loss of resolution and this loss can become interesting in itself.
The work also uses "spectral freezing" of components of the field recording to create spectrally derived chords from features of the recording bird sounds and a rusty gate which are then transcribed into notation for the instrumentalists and temporal manipulation of the recording to allow complex bird calls to be emulated in a human time-scale.
Nature Forms II explores the notion of eco-structuralism, maintaining what Opie and Brown (2006) term the "primary rules" of "environmentally-based musical composition": that "structures must be derived from natural sound sources" and that "structural data must remain in series".
The structure of the original work is conserved using the approach discussed in *the miracle of the rose"*, where the temporal proportionality of the recording is retained by aligning multiple notation and resynthesis versions of the recording in visual representations that can be alternated or combined in the creation process of the score, processing and fixed media (Figure 15.)

4. CONCLUSION

The processes described in this paper constitute a set of possible approaches to engaging with field recordings through machine and performative means. They provide a methodology for manipulating a "found sound" in a relatively precise manner through spectral analysis and synchronisation of visual and sonic elements of the work. 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 [11].
The imperfections in the transcription processes involved here are intriguing in themselves as they throw into relief

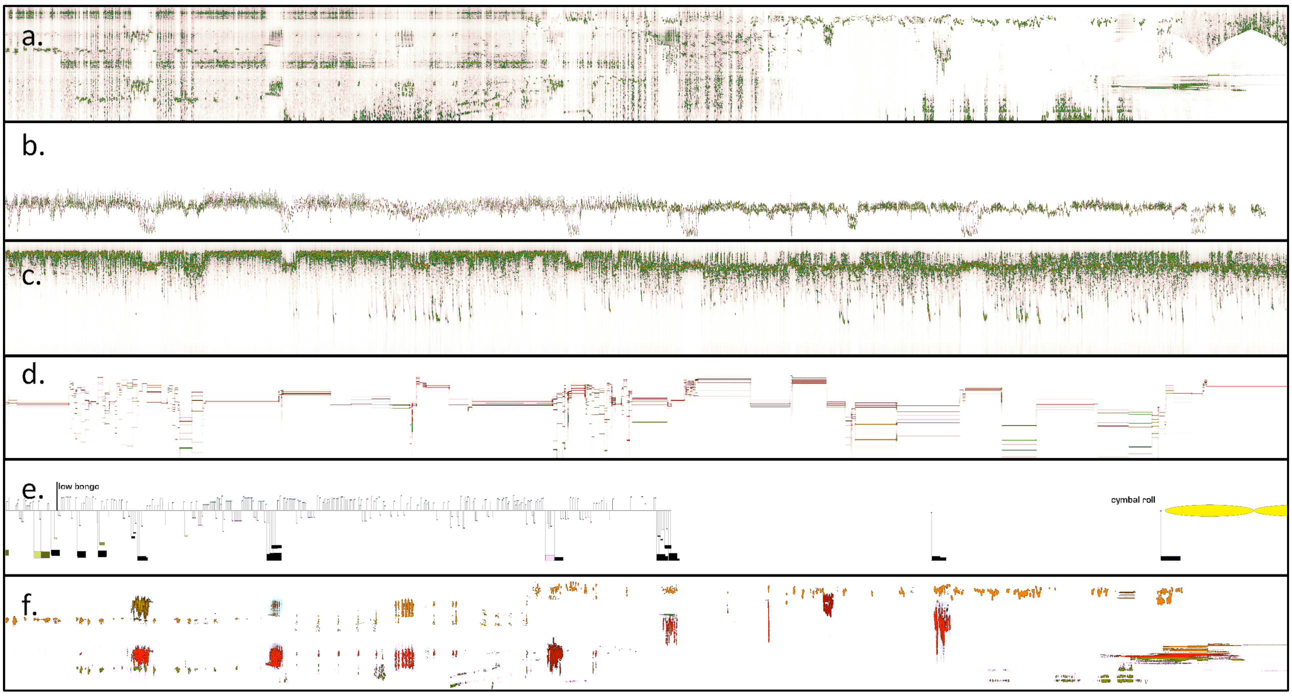

Figure 15. Visual representation of temporally proportional alignment of multiple resynthesis (a.-d.) and notation (e. – f) versions of the recording in *Nature Forms II* (excerpt): a. field recording spectrogram; b. ring modulation resythesis spectrogram; c. subtractive synthesis spectrogram; d. spectral "freeze" sonogram/score e *Nature Forms II* percussion score; and f. *Nature Forms II* instrumental score.

the distinctions between the various forms of representation: highlighting the gaps between "mimetic resynthesis" at different levels of abstraction. The implied circularity of processes opens the potential for re-interrogation of materials through the continuous transmutation of transcription by substituting source recordings for their resynthesised counterparts or live instruments and re-processing them.

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).

5. REFERENCES

[1] 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.

[2] Vickery, L., (2014a), Notational Semantics in Music Visualisation and Notation. *Proceedings of the Australasian Computer Music Conference 2014*, 101-112.

[3] Vickery, L., (2014c), The Limitations of Representing Sound and Notation on Screen. *Organised Sound*, 19(3), 215-227, Cambridge, UK

[4] Martino, G., and Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, 10, 61–65.

[5] Marks, L. E., and Odgar, E. C. (2005). Developmental constraints on theories of synesthesia. *Synesthesia; Perspectives from cognitive neuroscience*. New York: Oxford University Press. 214–236

[6] Jehan, T. (2004). Event-synchronous music analysis/synthesis. In Proc. Digital Audio Effects Workshop (DAFx), Naples, Italy.

[7] Harker, Alexander (2011) *AHarker Externals.* Alexander J. Harker. (Unpublished)

[8] O'Callaghan, J. (2015). Mimetic Instrumental Resynthesis. Organised Sound (20:2), 231 - 240

[9] Eimert, H. (1955). *Die Reihe* 1. Theodor Presser: Bryn Mawr: 4.

[10] Barrett, D. G. (2007). Music and Its Others. http://ablinger.mur.at/docs/barrett_others.pdf

[11] 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.