The ecological approach to mixing audio: agency, activity and environment in the process of audio staging

The Ecological Approach To Mixing Audio: Agency, Activity And Environment In The Process Of Audio Staging

Andrew Bourbon, Simon Zagorski-Thomas

Introduction

This paper reports on some of the outcomes from a larger research project on Classical Music 'Hyper-Production' And Practice As Research – a UK Arts and Humanities Research Council funded project that seeks to create radical reinterpretations of the classical repertoire through record production.

Our approach to mixing audio for this project is based on a theoretical model that explores the links between the perception and cognition of recorded music, our musicological analyses of the pieces and how that translates into staging and processing decisions. While taking into account Schaeffer's theories about the 'Objet Sonore' (Schaeffer 1966 and Dack 2001) and Smalley's (1986 and 1997) work on spectromorphology, we are utilizing the ecological approach to perception (Gibson 1979; Clarke 2005; Zagorski-Thomas 2014) and the neural theory of language and metaphor (Lakoff & Johnson 2003; Feldman 2008) to examine mix decisions in terms of agency, activity and environment.

Examples from the research project, which include ensemble pieces and layered, overdubbed solo performances, will be deconstructed from a musicological perspective. This will involve foreground and background, thematic material, contrapuntal lines and other musical features being discussed in terms of the number and type of perceived agents, the types of activity in which they are involved and the nature of the environment within which the activity occurs. This will be explored through both literal and metaphorical interpretations of the musical activity. These analyses will then be used to explain the decisions that were made during the mix process. Placing the perceived agents on different parts of the sound stage, highlighting or inhibiting various aspects of the energy expenditure involved in the perceived activity and determining the type and character of the environment within which this activity occurs will be further deconstructed in terms of the specific processing decisions that were made in different instances.

Theoretical Background

The main theoretical basis for this article is the ecological approach to perception. By this we mean active perception within our environment based on searching for invariant properties in the multi-modal perceptual field and identifying the affordances. In order to clarify, each phrase of this definition will be elucidated upon. "Active perception within our environment" entails two important features. By active perception we mean the purposeful direction of our senses and our cognitive activity to further explore features that come to our attention. In the case of hearing, this can involve moving one's head to aid with directional perception or to increase clarity by directing an ear towards a sound source so that we perceive more direct and less reflected or reverberant sound. We're more aware of our eyes continually moving to seek out detail and of using our fingers to explore the tactile world, but hearing is a similarly active and reactive process. The study of perception doesn't usually take this dynamic, time-based characteristic into account but it is hugely important and lies at the heart of Gibson's (1979) approach. Indeed, it is the 'flattening' of our environment in representational systems such as film, photography and stereo sound that is the primary mechanism by which we avoid being 'tricked' into thinking of them as 'real'. With stereo recorded sound, we either notice that the relative positions of instruments don't move in a 'natural' (i.e. three dimensional) way when we move in relation to a pair of speakers or, in headphones, we notice that they don't change position at all when we move.

Another important definition is of the "multi-modal perceptual field". Studies of perception have tended to examine individual modes separately i.e. sight, hearing, touch etc. Our perceptual system, however, works multi-modally at a very low level of cognition (see for example McGurk & MacDonald 1976): we are exploring the question of 'what is happening?' rather than 'what

am I seeing?' or 'what am I hearing?'. The McGurk effect demonstrates very powerfully – if you hear the sound 'ma' while looking at a face saying 'ba' you hear 'ba' and if you see 'fa' you hear 'fa'. Your perceptual system resolves the conflict at a pre-conscious level and even when you know it's happening you can't 'over-ride' it. This characteristic of the system, the resolution of perceptual conflict, also works within a single mode of perception. Optical illusions, for example, demonstrate both our need to resolve such conflicts on a micro-level, in that we can only 'see' one interpretation at a time, and our ability to flip between different interpretations (as in the 'vase / faces' illusion in figure 1).

Figure 1: 'Vase / Faces' illusion (**www.moillusions.com**)

This characteristic was crucial in allowing us to develop representational systems such as drawing and sculpture. Without this physiological 'need' to resolve perceptual information into an interpretation, we wouldn't have developed the ability to recognize a two dimensional visual outline of, say, a horse. A drawn horse provides certain patterns of light on the retina that are sufficiently similar to the patterns of light produced by the sight of a real horse to cause us to resolve the incomplete perceptual data into a horse-like interpretation. Of course, we're not tricked into believing we are seeing a real horse because we also have a parallel interpretation of, for example, a piece of paper.

The final pair of characteristics are "invariant properties" and "affordances". These two ideas work on a variety of levels. At the very basic level, the patterns of light on the retina are the invariant properties that afford a particular interpretation. The way that invariant properties suggest affordances is through experiential learning. For example, when we learn about bumping into things as an infant, the invariant properties of our embodied sense of propulsion becomes associated with the visual stimulus of light patterns moving from the centre of the retina to the periphery (i.e. the visual experience of forward movement). Some types of patterns on the retina become associated with a sudden stop and the sense of pain. We learn that particular invariant properties have various likelihoods of particular affordances. These connections are made at the very basic level of perceptual input – for example, the connection of visual perception of different sizes of space becoming associated with different lengths of reverberation. They also are made at higher levels of cognition – for example, developing a tacit understanding of musical structure so that we develop expectations about when we think a chorus might be approaching in a pop song.

At the neural level, when a set of neurons are triggered by a perceptual stimulus, the pathways that are triggered frequently become reinforced, i.e. where stimulus A is often followed by stimulus B, and we therefore develop predictive expectations, i.e. when we experience stimulus A we come to expect stimulus B. These develop into networks that function as what Lakoff (1990) and others called schemata: cognitive structures that embody our expectations about a particular object (an image schema) or activity (an event schema). A schema doesn't exist as an identifiable part of the brain but as a series of more or less likely pathways through the brain. Thus an image schema for a dog will be built from our previous experience of the visual, aural and tactile perception of dogs and a set of expectations about the kinds of actions they perform and the impact that has on the world and us. Our event schema for walking a dog will be built from a similar set of expectations. Of course, the distinction between image schemata and event schemata isn't one that exists in the brain but is just a useful category developed to help us describe brain activity. Indeed, the idea of a schema as a structure is also merely a useful categorical distinction rather than a physical phenomenon.

Schaeffer's 'objet sonore' (1966) can be seen as equivalent to an image schema: a cognitive construction that represents the abstract qualities of the sound rather than a single perceptive perspective on it. Schaeffer utilises Husserl's approach to phenomenology: a particular recording (or a processed playback) is an *adumbrage* – an incomplete view of the 'object' from a single perspective. The sound object has invariant properties that can be examined by 'looking' from a variety of perspectives and can be further examined through Husserl's notion of hypothetical substitution.

Smalley's notion of spectromorphology (1997; 1986) relates to the way that the frequency content of a sound changes over time. This morphology suggests a possible interpretation based on our experience of what kinds of agents, tools or objects, activity and environment create what kinds of noise. This relates very clearly to the ecological approach to perception: the holistic perception of what is happening in our environment. Smalley deals with 'unrealistic' and electronic sounds by examining how their connection with the 'cause' can be theorized in terms of levels of surrogacy: how similar they are to sounds we have experienced. Zagorski-Thomas (2016a; 2016b) has examined this idea of distorted, simplified or synthesized versions of 'real' sounds as *sonic cartoons* and has explored how some of the seemingly more abstract sounds of electronic music can be understood and interpreted in this way as well.

While the invariant properties of the sound are a result of the actors (human and non-human), activity and environment that produced it, Schaeffer's notion of reduced listening – listening to the properties of the sound and not to the causes – may seem incompatible with the ecological approach. However, as we have noted, the ecological approach is also about directed or active perception. We may not be able to access the deep-seated connectivity between sound and meaning that is created by our apprehension of cause and some metaphorical connection with embodied experience – but we can focus on particular properties so that we make new metaphorical connections. That type of re-direction of focus provides a theoretical basis for reduced listening in the ecological / embodied approach.

As Schaeffer pointed out, exploring the nature and possibilities of a sound object by recording and processing it in different ways was a key part of the *musique concrete* agenda and allowed for a creative exploration of the phenomenology and interpretative affordances of a sound object. Reduced listening afforded different directions in interpretation and therefore allowed composers to play with the invariant properties that encouraged those different directions This relates to Zagorski-Thomas' *sonic cartoons* in terms of the ways that the manipulation and distortion of particular invariant properties encourages particular affordances for interpretation. In extreme cases, the manipulation of a particular invariant property (or several) can lead to abstract, i.e. highly unrelated, interpretations.

An interesting phenomenon of this project was the way in which a study of the notation and the performance were important in the choices that led to the recording of the 'objets sonores'. The three examples that will be examined demonstrate how decisions in the recording process were designed to produce specific affordances for the mixing that had musicological reasons based on score analysis, the performance possibilities and listening-based analysis. The affordances that this process produced, however, while based on some specific ideas about interpretation (e.g. the metaphors of sea and mist in the Debussy), also gave scope for practical experimentation at the mix stage – the type of playing with invariant properties (by changing or distorting them) to encourage new forms of interpretation that Schaeffer engaged in with *musique concrete*.

Applying The Theory To Mixing Audio

To now move the discussion along to the question of how this theoretical framework can be applied to mixing audio, the key concepts are agency, activity and environment. In keeping with Latour's (2005) ideas from Actor Network Theory (ANT) that the actors (or agents) can be human or non-human, there is no separate category for tools or objects being used by the human actors to make noises. This distinction in ANT has caused considerable discussion about the nature of agency but, to simplify and shorten the argument, we are distinguishing between two types of non-human objects used to make music: those that are designed and those that are incidental or accidental 'participants'. Those that have been designed can be seen as the embodiment of the designer's agency and those that are incidental or accidental participants can be seen as part of the environment. On a very basic level then, this approach involves thinking about mixing in terms of influencing various aspects of how you hear 'who is doing what?', 'with what?', 'in which way?' and 'where?'.

The 'who' can be thought of primarily in terms of 'how many?' and 'what kinds?' of people. This may seem very straightforward and not something that can be influenced by the process of audio mixing: we can recognize the sound of one, two or three people quite easily and, beyond that, can make approximations about how many or few we might be hearing. We can recognize men, women and children quite easily and there are various cultural stereotypes that we may have experience of: the strong female gospel voice, the thinner Bollywood female voice, the deep Russian male voice choir, the western operatic voices. However, there are several subtleties that can be discussed here:

- Doubling a track in unison, particularly a vocal that is doubled accurately by the same person, can create the sense of a 'special' or thickened single voice rather than multiple agents (Warner 2005). This perception can be enhanced by treating the multiple agents with the same processing and effects. This 'fusion' is based on the Gestalt principle of common fate (Deutsch 1998).
- One of the key signifiers of the type of person we are hearing is the pitch of the voice. It's an invariant property of the world that larger objects make lower pitched sounds than smaller objects. One of the earliest techniques that was used in record production to alter our perception of the size of the singer was the Chipmunks Song (The Chipmunks 1958) where three of the four voices were recorded at half speed and then sped up to create high pitched voices that represented small cartoon chipmunks.

- A similar feature is the weight or resonance of a voice. By removing some of the resonance with equalization or filtering, the size of the singing body can be made to seem smaller.

The 'doing what?' aspect also seems at first glance to be something that would not be alterable by audio processing. It's hard to imagine that the mix process could change our perception of, say, a bowed sound to a plucked sound. For the most part that is true but there are many instances of sounds being combined together to disguise one kind of activity as another. For example, the addition of percussion sounds to a vocal track can disguise the percussive sound as a plosive vocal sound by putting it in a new context. An example of this is the scraped guiro sound that producer Max Martin mixed with Britney Spears' vocal fry on 'Oops!... I Did It Again' (Spears 2000). The guiro becomes perceptually fused with the voice, ceasing to be the sound of stick scraping on a ribbed piece of wood and, instead, becoming the grating sound of a constricted throat.

The 'with what?' is an interesting phenomenon because several techniques that started out as studio-based processing have led to changes in the design of physical objects. As equalization and noise gates were used more on drum sounds in the 1970s, they changed the perception of what type of objects were hitting and being hit. In order for drummers to be able to duplicate those sounds in live performance, drum manufacturers changed the materials and the design of the drums: adding nylon tips to drum sticks to emulate the additional high frequency added through equalization and creating various material changes to drum skins to emulate the shorter sounds that noise gates created (Zagorski-Thomas 2010). The processing had so altered the perceived materiality of the instrument that manufacturers redesigned the instruments to emulate that alteration.

The 'in which way?' can relate to the perceived type and level of energy involved in the activity that made the sound and this seeps into almost every other aspect of the process. Thus, while Max Martin's trick can be seen to have changed the nature of what was perceived as being done with the guiro, the main effect was to change the perception of how Britney was singing i.e. it added a perception of tension and angst to the vocal performance. In the same way, equalizing some of the weight out of a performance alters the perception of the amount of energy being expended on the activity as much as it alters the perception of the size and weight of the agent. The perception of energy can also be altered by adding distortion or overdrive to a performance: guitars are the most ubiquitous example of this but it's used on drums, vocals and other sounds as well to create a slight energetic 'bite'. Parallel compression of various forms also alters the perceived level of energy expenditure.

The 'where?' is arguably the oldest and most firmly established form of signal processing because it can be altered by microphone technique as well as the architecture of the recording space. The concept of echo chambers and, later, of artificial reverberation extended the possibilities even further and also created the prospect of 'impossible' and 'abstract' spatial recordings. Even Miles Davis' iconic *Kind Of Blue* (1959), while seeming to be quite a naturalistic recording, utilizes different reverberation on the trumpet and two saxophones than on the piano, drums and bass by using an echo chamber. It was on these types of recording that techniques which distorted 'reality' in order to create greater clarity were developed: ideas such as filtering out the low frequency rumble from an echo chamber so that you got the sense of space without the muddiness or using a delay on the reverberation so that you heard the direct vocal clearly before the reverberation starts to sound were developed in the 1950s and 1960s.

This theoretical approach, therefore, provides both a way of analyzing the things that have been done with record production and a way of thinking about what you might do. The latter provides a way of thinking about the process from the perspective of the musical sound rather than the techniques and that makes the kind of lateral jump that is involved in innovation more intuitive. Thus, thinking about what you want to add or alter in terms of the energy of a particular musical element allows you to think about the various invariant properties that afford that type of interpretation. That way, you can think about the range of types of processing that would produce those affordances, which properties might be undesirable and how you might avoid or inhibit them. On the other hand, thinking in terms of techniques – 'what would I (or one of my heroes) normally do in this kind of situation?' – tends to lead to fixed ways of thinking and a lack of flexibility.

Debussy Example

This first example, Debussy's piano prelude, La Cathèdral Engloutie (the Sunken Cathedral) is the subject of another paper in this issue of the journal (Capulet & Zagorski-Thomas 2016) and will, thus, be dealt with in less detail than the other two examples in this article. Three parameters that were important in this piece were volume, intensity and distance. We were playing with various invariant properties that encourage particular interpretations about these three parameters. One example of this is the 'Bells' that were suggested by a repeated sustaining single note were separated from the rest of the performance through overdubbing and played loudly but then staged to be quiet and distant (hear audio example one).

00:00   00:00

Audio example one

In another aspect, we were looking for metaphors for sea and mist and used the idea of reverberation as a reducer of clarity as well as a creator of space. This allowed us to expand on the idea by taking Debussy's programmatic description of the

enchanted sunken cathedral rising from the sea and mist. We utilized a metaphorical connection between invariant properties here by representing this phenomenon through a changing ratio of direct and reverberant sound: the direct sound being the object and the reverberation being the mist or sea. The mist therefore recedes and the object approaches.

Mozart Example

The next example is a recording of Mozart's string quartet no.19 (KV465) performed by Malgorzata Filipowicz (vln 1), Anna Olszewska (vln 2), Kamila Barteczko (vla) and Anna Kulak (vcl) and arranged by Simon Zagorski-Thomas. It was recorded to a click track in a series of different spaces in a 16th century palace at Piotrowice Nyskie in southern Poland in June 2015 (see video 1).

Video Example 1 – Mozart-Rooms

The arrangement involved changing very little musically but fragmented the parts so that they could be recorded in different spaces [1]. The inspiration for this was an analysis of the Kings Of Leon's 'Sex On Fire' (Zagorski-Thomas 2015) and, in particular, the use of different acoustics in various sections of that production that roughly corresponded to the notion that louder passages were staged in larger spaces. Relating this back to the ecological approach to perception, there is a definite metaphorical connection between the ideas of energy and size. We expect activity in a smaller and more intimate space to be conducted at a lower energy level than activity in a large and more public or impersonal space. Of course, what we wouldn't expect in 'reality' is for the size of the space to change with the energy level of the activity but it makes sense as a metaphor. With this as a guiding principle, the quartet was arranged so that parts could either move suddenly into different spaces or could cross fade between spaces during crescendo and diminuendo passages. Another production decision was to record the eighth note staccato parts in a way that would allow them to be mixed like a palm muted guitar 'chug'. That required they be as 'dry' as possible and so they were recorded very close and screened off from the small room they were in with heavy curtains and blankets.

One of the key features of this recording, then, was that it was planned and designed with a very specific mixing strategy in mind. Looking back on the recordings it is really easy to see a visual roadmap through the piece simply by looking at the positioning of waveforms in the session. The session is arranged according to the different spaces and you can see that mapping visually (see figure 2). The seven spaces that were used were:

1. The master bedroom with a domed ceiling. This provided a small chamber reverb and the majority of the recording was done one instrument at a time.
2. A medium sized barn space where the recording was done one instrument at a time with a close microphone and another about four meters above the player in the roof.
3. A corner of a small living room with each player recorded individually with a close microphone and heavy curtains and blankets to deaden the acoustic.
4. A small basement with a low ceiling and stone and brick walls and floor. The quartet played together because the dampness in the basement made it difficult to work there for long periods of time.
5. A cow shed where the ceiling consisted of a series of parabolic domes that created very intense comb filtering and resonant reverberation. Each player was recorded separately.
6. A large barn where the recording was done one instrument at a time with a close microphone and another about thirty meters away.
7. The ballroom where a short segment was recorded with the quartet playing together.

The contrast between lower dynamics in smaller spaces and louder parts in larger spaces was made more interesting by having two contrasting pairs of small and large ambience. Thus spaces one and two provided one pairing (with space six providing a more extreme version of the large space) and spaces four and five provided a second pair that provided more resonant and intense versions of small and large. It should be mentioned that an additional factor that ensured that the recordings would have to remain an experiment rather than a finished recording was the noise. The palace is home to dozens of poultry – from hens and geese to a turkey – and the recordings are infused with the sound of the birds. Andrew Bourbon spent a lot of time and energy, mostly in post-production, editing the multi-tracks to minimise the presence of the livestock but it was always clear that this noise would make it impossible to record studio quality masters.

These experiments did, however, flag up two very important and interesting issues that arose in relation to the spaces and the noise. The first of these was the effects that were caused where parts exist in multiple spaces and were mixed together e.g. listening back to the same part playing in spaces one and two simultaneously. While this adds a complexity and richness to the sound of the ambience, it does so in a schematic way: making the sensation of space less real and more abstract. The reverberation seems to become almost two dimensional and the layering of time differences and the un-natural summing of spatial cues adds an odd density to the ambience, sounding like more players at times and in other examples simply sounding bigger and more complex. The use of multiple recordings in different spaces at the same time hadn't been part of the pre-production plan and it was only during the listening process in preparation for mixing that the richness of this phenomenon

became apparent. However, as the recording progressed and we realised that the 'bird noise' was going to make a final, professional level mix unlikely, we also realised that there were many possible mix experiments inherent in the recordings we were making.

00:00 00:00

Audio example two

At time of writing there are two main mixes in existence and a third one in preparation. The first is a version embracing the spatial possibilities that were pre-planned and later discovered in the production process (hear audio example two). We used some of the initial plan but also explored the potential of layering spaces in certain situations. In this mix the main tasks were technical issues with variable noise floor and performance differences. The noise floor is an incredibly important part of this process as, leaving aside the extraneous livestock noise, the space itself exhibits noise that is detected by the listener and can significantly reduce the musical impact of transitions. The preparation process for this mix involved working quite intensively on the ways that these cross fades between recordings in different spaces might take place. This involved experimenting with different invariant properties of the spatial states, including their characteristic noise floors, to discover a 'choreography' of transition which produced a sense of movement between spaces without allowing the change between noise floors to distract attention away from the music. This does highlight one of the problems that has faced producers of popular music in various forms, that the creative use of production techniques can draw listeners' attention towards features other than performance and composition. Even though the production strategy was to build the spatial movement around musical changes so that they become part of that musical change, we are coming up against the *habitus* (Bourdieu 1993) of listening. The aesthetic of the concert hall sound is so firmly entrenched in classical music recording (as it also is in most jazz and folk recording), that it is hard to move far away from that aesthetic without it becoming distracting. The various styles of recorded popular music took several decades to develop and this was accompanied by a constant narrative about 'over-production' and distracting from the music (Frith 2012). Although we undertook this research project with the explicit intent of experimenting with, and pushing at the boundaries of, listener expectations in recorded classical music, it has been interesting to note how we have had to establish our own criteria for judging what is musically appropriate: about what works and what doesn't.

00:00 00:00

Audio example three

The second mix (hear audio example three) starts to take this idea further by embracing the previously discussed functional elements in the arrangement and looking to enhance those functions through processing usually reserved for band based western popular music. The rhythmic 'chug' of the staccato parts, for example, was processed using techniques often employed for layered rhythm guitar parts. The invariant properties being explored in this case are to do with the rhythmic momentum and the controlled use of energy. In addition, lead lines were stabilized with compression and brought to the front and centre of the image, bypassing the established panorama associated with a traditional string quartet. The existing spaces are occasionally enhanced through artificial reverb, with distortion and time-based effects used to draw attention to and manipulate the impact and energy of targeted parts.

The third mix, which at the time of writing is not complete, will eventually be available on the University of West London website (**http://www.uwl.ac.uk/classical-music-hyper-production/welcome**). Having explored some of the creative possibilities of the spatial recordings in the first two mixes, it became apparent that the layered recordings were sounding like much larger ensembles than a quartet (this was especially true of the second mix). This is another of the issues that can be found within the history of popular music production: a band augmenting their line up in the recording studio to the extent that the identity of the original ensemble can become blurred or lost. This approach, then, aims to strip back the arrangement and to return to the sense of a quartet rather than a string ensemble.

Palestrina Example

The third example is a recording of a fragment of Palestrina's motet *Ad Dominum Cum Tribulare,* performed by Katarzyna Stern (soprano), Angelina Roch-Domagala (alto), Marek Murawa (tenor) and Lukasz Smolka (bass). This was also recorded in Piotrowice Nyskie in the same week as the Mozart quartet and involved some small additions to arrangement by Simon Zagorski-Thomas. The inspiration for this recording strategy was Simon and Garfunkel's 'Bridge Over Troubled Water' (1970) which starts out very small and intimate and grows throughout the arrangement into something more large scale and epic. We were also interested in exploring ways of creating stereo effects in the recording process. This involved recording the whole piece to a click track six times:

1. The small acoustically 'dead' living room that was used in the string quartet recording. Each voice was recorded separately.
2. A small but four metre tall room in the palace's tower. Each voice was recorded separately.
3. The small basement with a low ceiling that was used in the string quartet recording. All voices were recorded together.

4. A dining room with stone walls and a sprung wooden floor that was about 40m^2. A mid-side microphone array was set up in the middle of the room and the singers walked slowly around it in a circle as they performed as an ensemble.
5. The cow shed with the parabolic domes that was used in the string quartet recording. All voices were recorded together.
6. The large barn that was used in the string quartet recording. In this instance, two omni microphones were slung over the roof beams on either side of the ensemble who sang together. The microphones were swung like pendulums during the recording to create a swirling stereo effect.

The Palestrina recording explored similar characteristics to the Mozart, using space and layering to create expansions in size and also in the perceived number of 'agents' in the space. The fundamental difference in the approach taken here is that rather than pre-scoring into the available spaces the piece was performed in full in multiple spaces, allowing the piece to be presented from any of the spaces at any point in the work. The performers were recorded together in the relatively controlled environment of space one to create a guide track on which the other performances could be built. They were then recorded individually in the same space, giving the ability to offer focus on any part at any time. The contrast between this very dry space, full of detail and proximity, provided a very useful tool for creating contrast to the other spaces and allowing us to mix the timbral detail of that direct sound into different spaces. This technique of placing an intimate recording in a larger space was used on 'Bridge Over Troubled Water' through the use of an echo chamber and is, of course, a staple method for maintaining the invariant properties of intimate communication (i.e. proximity and a relatively low intensity of performance). As the other available spaces were explored the ensemble was recorded together in some spaces, and then individually in others, providing the ability for contrast and also allowing exploration of some unusual spaces, including a space with the longest dimension being the height and room for only a single performer and engineer on the floor (space 2). The final space in which we recorded was the large barn, with the microphones rigged from height on long cables, which were then put into motion as opposing pendulums. Though some modification was required to remove any wind and physical noise these microphones provided some form of movement through destabilization of the stereo image. The technique provided something subtler than dramatic panning changes, instead creating a feeling of texture and life through this lack of stability.

00:00   00:00

Audio example four

Once recorded the mixing process differed greatly from those employed in the Mozart. The plan for this piece was to have a feeling of spaces expanding throughout the performance, opening virtual doors into the different spaces captured in the recording. The mix strategy involved a 'collage' approach to the recordings, exposing parts during the piece through muting tracks to move between spaces throughout the performance (hear audio example four). Initially the piece starts with the smallest of the spaces, using individually recorded lines in spaces as triggers for spatial development. In the second half of the piece the listener hears a second development, with the addition of multiple layers of spaces to the initial simple spatial expansion. The feeling of increased performer numbers adds depth and excitement, but also has quite a unique impact created by the layering of four voices into what can be perceived as a choir by the end of the work. While multi-tracking a single voice sounds very different to recording a choir because of the single timbre, multi-tracking four voices does seem to create more of a choir effect even if each voice multi-tracks a single line.

Discussion And Conclusions

As mentioned earlier, the discussion of invariant properties and affordances in relation to mixing audio can be looked at in terms of the number and type of perceived agents, the types of activity in which they are involved and the nature of the environment within which the activity occurs. Mixing strategies have been developed for the three examples discussed here which deal with each of these aspects in both literal and metaphorical terms. In the Mozart example, the decisions had strong implications about the perceived number of agents in a very literal sense and the mix strategy was based on the literal narrative of relating performance intensity to the size of the physical space. In the Debussy example, the entire narrative structure was build on a metaphorical interpretation of the piece drawn from the mythological story that Debussy used as inspiration for the composition.

Both Smalley and Schaeffer have been assimilated into a theoretical framework based on the ecological approach to perception and the neural theory of language and metaphor. Smalley's *spectromorphology* relates directly to the notion of perception and interpretation being built upon our embodied experience and an ability to make metaphorical connections between new and previous experience: a more nuanced version of Smalley's surrogacy. Schaeffer's *objet sonore* relates similarly to the notion of image schemata. Both have been related to Zagorski-Thomas' *sonic cartoons* and the implications that this way of thinking about recorded sound has for approaches to mixing in general as well as this project in particular.

Bibliography

Bourdieu, P. (1993) *The Field of Cultural Production*, New York: Columbia University Press.

Capulet, E. & Zagorski-Thomas, S. (2016) Creating A Rubato Layer Cake: performing and producing overdubs with expressive timing on a classical recording for 'solo' piano. *Journal on the Art of Record Production*, (11). Available at: **http://www.arpjournal.com/creating-a-rubato-layer-cake**.

Clarke, E.F. (2005) *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*, Oxford University Press, USA.

Dack, J. (2001) At the limits of Schaeffer's TARTYP. In *Nowalls Conference Proceedings*. Nowalls Conference. De Montfort University: De Montfort University Research Repository. Available at: **http://www.dmu.ac.uk/documents/technology-documents/research/mtirc/nowalls/mww-dack.pdf** [Accessed March 28, 2015].

Davis, M. (1959) *Kind of Blue*, Columbia CS8163.

Deutsch, D. (1998) *The Psychology of Music* 2nd ed., Academic Press.

Feldman, J.A. (2008) *From Molecule to Metaphor: A Neural Theory of Language* 1st MIT Press Paperback Ed., MIT Press.

Frith, S. (2012) The Place of the Producer in the Discourse of Rock. In S. Frith & S. Zagorski-Thomas, eds. *The Art of Record Production: an introductory reader to a new academic field*. Farnham: Ashgate Publishing Limited.

Gibson, J.J. (1979) *The Ecological Approach to Visual Perception*, Psychology Press.

Lakoff, G. (1990) *Women, Fire, and dangerous Things: What Categories Reveal about the Mind*, Chicago: University Of Chicago Press.

Lakoff, G. & Johnson, M. (2003) *Metaphors We Live By* 2nd ed., University Of Chicago Press.

Latour, B. (2005) *Reassembling The Social: and introduction to Actor Network Theory*, New York: Oxford University Press.

McGurk, H. & MacDonald, J. (1976) Hearing lips and seeing voices. *Nature*, 264(5588), pp.746–748.

Paul Simon & Garfunkel, A. (1970) *Bridge Over Troubled Water*, Columbia KCS9914.

Schaeffer, P. (1966) *Traité des objets musicaux*, Paris: Le Seuil.

Smalley, D. (1986) Spectromorphology and Structuring Processes. In S. Emmerson, ed. *The Language of Electroacoustic Music*. London: Macmillan, pp. 61–93.

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

Spears, B. (2000) *Oops!…I Did It Again*, Jive 9250582.

The Chipmunks (1958) *The Chipmunk Song*, Liberty F-55168.

Warner, T. (2005) The Song of the Hydra: multiple lead vocals in modern pop music recordings. In Art of Record Production Conference. Westminster University. Available at: **http://www.artofrecordproduction.com**.

Zagorski-Thomas, S. (2015) An Analysis of Space, Gesture and Interaction in Kings of Leon's 'Sex On Fire' (2008). In R. von Appen et al., eds. *Twenty-First-Century Pop Music Analyses: Methods, Models, Debates*. Farnham: Ashgate Publishing Limited.

Zagorski-Thomas, S. (2010) Real And Unreal Performances. In A. Danielsen, ed. *Rhythm In The Age of Digital Reproduction*. Ashgate, pp. 195–212.

Zagorski-Thomas, S. (2016a) Sonic Cartoons. In M. Haná?ek, H. Schulze, & J. Papenburg, eds. *Sound As Popular Culture*. Cambridge, MA: MIT Press.

Zagorski-Thomas, S. (2014) *The Musicology of Record Production*, Cambridge: Cambridge University Press.

Zagorski-Thomas, S. (2016b) The Spectromorphology Of Recorded Popular Music: the shaping of sonic cartoons through record production. In R. Fink, M. L. O'Brien, & Z. Wallmark, eds. *The Relentless Pursuit Of Tone: Timbre In Popular Music*. New York: Oxford University Press, USA.

Notes:

[1] The only changes were to some of the staccato rhythmic parts to fill out the harmony.