Electroacoustic Music with Moving Images: the art of media pairing

JOHN COULTER

C/- School of Music, National Institute of Creative Arts and Industries, University of Auckland, 6 Symonds, Auckland, New Zealand
E-mail: j.coulter@auckland.ac.nz

Composers working with sounds and moving images are immediately confronted with a paradox. On one hand, audiovisual materials appear to offer the possibility of complementing one another – of forming a highly effective means of communicating artistic ideas – and on the other, they appear to carry the risk of detracting from one another – of deforming the musical language that he or she has worked so hard to create. Durk Talsma and Max Mathews succinctly state the opposing principles. 'Many behavioural studies have provided evidence for the hypothesis that integrating visual and auditory stimuli serves the purpose of enhancing perceptual clarity … These results suggest that communication between the visual and auditory brain areas is a highly effective and relatively automatic process' (Talsma, Doty and Waldorff 2007: 679). 'I personally find most combined music-video art problematic. It seems to me that the sound and images often compete for my attention … If I pay attention to what I am seeing, I often miss what I am hearing, and if I try to concentrate on the music, the images can often be an irritating distraction…' (Mathews 2007: 94). This article seeks to transcend this paradox through the identification of audiovisual materials that function in different ways. Examples of creative work are offered to illustrate more general points of 'language', a model for classifying media pairs is put forward, and practical methods of audiovisual composition are proposed. The narrow findings of the study offer a vocabulary for discussing the functionality of audiovisual materials, detailed methods of media pairing, and techniques of parametric alignment, while the wider findings extend to associated domains such as live electronic music, and hyper-instrument design.

1. ACOUSMATIC AND AUDIOVISUAL MODES

There are a number of definitions of the term 'acousmatic' at our disposal. Leigh Landy encapsulates the essence of meaning as follows.

> Acousmatic listening is the opposite of direct listening, which is the 'natural' situation where sound sources are present and visible … By isolating the sound from the 'audiovisual complex' to which it initially belonged, it creates favourable conditions for reduced listening which concentrates on the sound for its own sake, as sound object, independently of its causes or meaning … By repeated listening to the same recorded sound fragment, the emphasis is placed on variations of listening … [which arise from] directions which are always precise and always reveal a new aspect of the object, towards which our attention is deliberately or unconsciously drawn. … Indeed if curiosity about causes remains in acousmatic listening (and it can even be aroused by the situation), the repetition of the recorded signal can perhaps 'exhaust' the curiosity and little by little impose 'the sound object' as a perception worthy of being listened to for itself, revealing all its richness to us. (Landy 2007: 78).

The audiovisual mode, in its widest sense, concerns all combinations of vision and sound; however, for the purposes of this discussion, we will limit the definition to include simultaneously attended combinations of onscreen moving images and electroacoustically produced sounds. Michel Chion describes the language of sound with moving images as being primarily founded on visual language, referring to 'cinema' as 'a place of images, plus sounds' with sound being 'that which seeks its place'. Here, and in similar settings, he proposes that listeners enter an 'audiovisual contract' where it is the function of sound to 'add value' to the visual images presented (Chion 1994: 68).

Differences between the acousmatic and audiovisual modes are evidenced by empirical studies – as Talsma states: 'It has been found, however, that auditory stimuli are known to capture attention easily (Schroger et al.) and also that the processing of audio stimulus features occurs generally faster than that of visual stimuli (Waldorff et al.) Thus, on the basis of these differences in processing time, we would also predict that attending to the visual modality would affect the multisensory integration process differently than attending to the auditory modality' (Talsma et al. 2007: 681).

There are two points to be raised concerning the relationship between 'the acousmatic' and 'the audiovisual' before moving to the main body of the study (which concerns only the audiovisual mode).

Firstly, it is clear that the two modes are closely aligned. For most, the acousmatic includes a quasi-visual component, while both real and imagined materials constitute the experience of the audiovisual mode. Depending on the receiver's disposition, there can be very little separating the two experiences. Simon Emmerson elaborates:

> The acousmatic condition excludes clear information on source and cause which we (products of evolution)

Organised Sound **15**(1): 26–34 © Cambridge University Press, 2010. doi:10.1017/S1355771809990239

attempt to 'fill in'. For me (and I believe many others) that process has a visual component. The imagination constructs a quasi-visual mindscape with many of the characteristics of 'real' vision. There are also interesting cross-references to other sense modes – most notably a sense that I might be able to 'feel' textures. (Emmerson 2007: 86).

Secondly, notwithstanding the similarities, transition between the two modes can be described, in essence, as 'a switch' – or more precisely as 'a tolerant switch'. Chion attests to the absence of a continuum between the two modes, noting that a collapse of 'the audiovisual structure' is necessary before a transition to the general audio-only mode can occur (Chion 1994: 40). In the writer's 2007 paper 'The Language of Electroacoustic Music with Moving Images' a series of tests were reported confirming Chion's views. As part of the study, the effects of 'black screen', memory residues, minimal images, and colours were investigated. It was found that 'black screen' images tend to evoke a condition similar to that of the acousmatic mode, while minimal moving images tend to demand audiovisual attention in the same manner as more complex moving images. A notable exception was when 'white screen' was used in place of 'black screen'. The majority of participants experienced sensations of 'retaining' the acousmatic mode; however, they also reported that the associated sounds took on 'white' characteristics: hence 'the tolerant switch' (Coulter 2007: 7).

2. A MODEL FOR CLASSIFYING MEDIA PAIRS

The 'dual potential of sounds [and images]' provides the foundation for the classification of media pairs. Denis Smalley illuminates the principle: 'All sounds possess this dual potential – the abstract and concrete aspects of sound – and all musical structures are balanced somewhere between the two, although exactly how they are found to be balanced can vary greatly among listeners' (Smalley 1986: 64). Decussation yields four groupings: referential video–referential audio, referential video–abstract audio, abstract video–referential audio, and abstract video–abstract audio.

In attending these audiovisual materials we naturally tend to integrate the component textures – as Francisco Kröpfl states (paraphrasing John Cage) 'even in extreme cases of heterogeneity between both dimensions mind needs to establish a coherent link in a process of simultaneity, thus, in spite of a non-existent will of organisation an structural integration is inevitable' (Kröpfl 2007: 90). However, experience tells us that we are also capable of separating the two components. This allows us to evaluate each media type independently, and, in remote circumstances, to attend the audio and visual components as dissociated

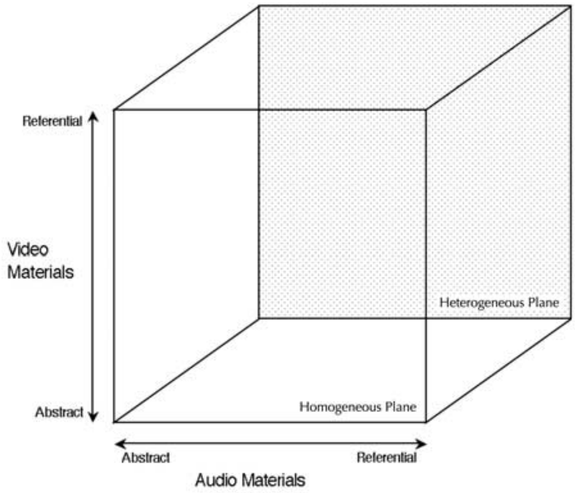

Figure 1. A model for classifying media pairs.

information stems. An example of this phenomenon is driving a car while talking on a cell phone. In situations that require such divided attention, our cognitive ability also tends to be divided. Hence the concern for road safety, and the phrase 'your full attention is required!'

The Necker Cube pictured in figure 1 provides an established construct for considering our natural tendency to integrate, and our ability to separate audiovisual materials. It is a psychological switch of sorts – one that allows us to oscillate between homogenous and heterogeneous experiences.

3. CONCOMITANT AND ISOMORPHIC RELATIONSHIPS

Integration occurs involuntarily when congruous relationships are perceived to exist between audio and video materials. Two types of relationships have been observed: 'concomitant', defined as 'going together, accompanying, concurrent, attendant', and 'isomorphic', defined as 'groups or other sets corresponding to each other in form, and in the nature and product of their operations' (OED 2009).

Media pairs exhibiting concomitant relationships are integrated through a process of highlighting and masking. Materials that support homogeneity automatically assume primacy over those that do not, resulting in a hierarchy of textures – the greatest of which have the potential to form the structure and meaning of the work, and the least of which are entirely ignored. Chion refers to this phenomenon as 'synchresis'. He states: 'Play a stream of random audio and visual events, and you will find that certain ones will come together through synchresis and other combinations will not' (Chion 1994: 65).

Media pairs displaying isomorphic relationships contain features that act as catalysts in the process of integration. These qualities include shared formal characteristics such as duration, intensity, proximity, and so on (abstractions), as well as shared meanings (references). When both form and meaning are aligned, we experience highlighting without masking. An example of this phenomenon is the experience of watching and listening to a newsreader (on a television) in a quiet and dimly lit room.

In essence, materials that exhibit concomitant relationships rely on the highlighting and masking that occurs when two (or more) schemas are simultaneously activated (overlaid), while materials exhibiting isomorphic relationships rely on the activation of solitary schemas. However, it is important to note that this binary description is an oversimplification. In reality, audiovisual materials exhibit instantaneous combinations of concomitant and isomorphic relationships, and exactly how the combinations affect the process of integration varies greatly amongst listeners/viewers.

In considering the two types of relationships in the context of the four classes of media pairs already identified, eight different possibilities are arrived at (one for each corner of the cube). The categories are abbreviated as follows: refV–refA–con, refV–absA–con, absV–absA–con, absV–absA–con, refV–refA–iso, refV–absA–iso, absV–absA–iso, absV–absA–iso. Examples of each class are presented and discussed in the following sections.

4. CONCOMITANT MEDIA PAIRS

Movie example 1 (an extract from an early version of the writer's 2005 work *Shifting Ground*) concerns referential video and referential audio displaying concomitant relationships. Common referential features are highlighted (in this case stones, the sun, etc.) while other qualities less relevant to the process of integration become masked (in this case, perhaps the image of the sky, or the spoken word 'angel').

Movie example 2 (an extract from *Shifting Ground*) concerns referential video and abstract audio displaying concomitant relationships. The category relies on the principle of 'primacy of reference' (in this case video-dominant) referred to by both Nicholas Cook and Durk Talsma. 'There is a general methodological point here. Whenever one medium appears to have a relationship of primacy over another – whether in terms of production or reception – *inversion* of the relationship becomes a useful heuristic procedure' (Cook 1998: 81). 'Audiovisual integration processes appear to associate the visual and auditory stimulus components with each other, even when only the visual component was relevant' (Talsma et al. 2007: 686). During the production process of *Shifting Ground*, a number of referential sounds were transformed and mixed (abstracted) to create the impression of rocks falling. When attended in isolation, several referential sounds may be clearly identified including: fireworks, marching band ('Snoopy's Christmas'), bagpipes ('Scotland the Brave'), crowd noise, cheers, Happy Birthday, a Maori haka, and individual voices; however, when accompanied by the moving images of rocks falling, almost all of these sounds are masked, and the majority of listeners/viewers hear only 'rock slides'.

Movie example 3 (another extract from *Shifting Ground*) concerns abstract video and referential audio displaying concomitant relationships. Once again, the category relies on the principle of 'primacy of reference' (in this case audio-dominant). Various components of the abstract video textures are highlighted (and given meaning) by specific words in the accompanying poem (such as shining, glistening, blue, etc.), while other qualities less relevant to the process of integration are masked.

Movie example 4 (another extract from *Shifting Ground*) concerns abstract video and abstract audio displaying concomitant relationships. This category relies on the intersection of abstract 'images' created by both the audio and video textures. In this case, visual images could be described as a remote representation of the acousmatic space (white, cloud-like, etc.), while the accompanying gestural sounds could be described as objects that exist within the space.

To summarise at this point, materials that exhibit concomitant relationships simultaneously activate (overlay) two or more (heterogeneous) schemas. As a result of the process of integration (towards the homogeneous plane), audio and video textures become highlighted and masked. The principle of 'primacy of reference' applies in all circumstances, although in situations where referential qualities are matched or undetermined primacy may oscillate between audio and video textures.

One more layer of complexity will now be added. Both audio and video materials also contain internal Newtonian relationships. The sound of breathing, for example, is inherently related to the human body, while the image of waves breaking has an inherent relationship to the shore. In considering these two schemas and their internal relationships independently, we arrive at the formulae A is to B (breathing is to the body) and X is to Y (waves are to the shore). In overlaying the two schemas, we arrive at the formula A is to B as X is to Y (breathing is to the body as waves are to the shore). This immediately highlights the common feature(s) (in this case the envelope) and masks the majority of other details. Aristotle first proposed this equation as a means of constructing metaphors, and as an illustration of the way in which we learn. The metaphor is produced by switching B and Y, resulting in the equation A is to Y as X is to B (In this case 'breathing is to the shore as waves are to the body') The resulting metaphors are 'breathing–shore' and 'wave–body' – both of which are in common use. As A–B and X–Y relationships become closer, more common aspects become highlighted. Conversely, as A–B and X–Y relationships become more remote, more tentative associations become highlighted. The example 'wind is to fire, as distance is to love' illustrates such a remote association. In this case the explanation 'it fans the great and extinguishes the weak' is normally required to identify the commonality. In applying Aristotle's equation, however, we produce 'wind–love' and 'distance–fire' – two remote, yet highly descriptive metaphors.

5. BUILDING CONCOMITANT RELATIONSHIPS

Choosing audio and video materials displaying concomitant relationships is a relatively intuitive process. Our involuntary tendency to integrate even the most remote of heterogeneous textures ensures that associations will be formed irrespective of our intentions. Identifying and controlling those relationships as means of artistic expression, is, however, somewhat more difficult. Aristotle's formula can provide a useful heuristic procedure for determining the accessibility of creative ideas articulated through the use of audiovisual materials displaying concomitant relationships (most relevant to the use of refV–refA–con materials). If commonly used metaphors are produced when contemplating the referential content of materials, then a simple association of audio and video media should be sufficient to transmit the creative idea. If more remote metaphors are produced, then additional information (perhaps even explanation) may be required. This empirical approach does not, however, remove the necessity of beginning with quality ideas – it merely provides a means of testing them.

6. ISOMORPHIC MEDIA PAIRS

Movie example 5 (a contraction of the writer's 2008 work *Mouth Piece*) concerns the use of referential video and referential audio displaying isomorphic relationships. The work makes use of an electronic device that transforms ordinary speech into musical tones. The piece begins with a spoken phrase, which is then repeated (hundreds of times). Sounds and images are gradually transformed, while the initial isomorphic relationships (both referential and abstract) are retained. The function of memory assists in housing the transformed sounds and images within a solitary schema that also contains the initial spoken phrase. I refer to this technique as 'schematic extension'.

Movie example 6 (an extract from an early version of *Mouth Piece*) concerns the use of referential video and abstract audio displaying isomorphic relationships. The category relies on the capability of referential video and abstract audio materials to form relationships that function as catalysts in the process of integration (towards the homogeneous plane). Here, the principle of 'primacy of reference' is at its most extreme, in that the abstract audio textures will normally be perceived as entirely subservient to the referential video materials. An example of this phenomenon is found in early Disney cartoons, where abstract instrumental sounds were used to represent the real-world actions of cartoon characters. A violin played in a particular way, for example, might imitate speech. This is in fact a form of mimesis (in this case the primary form would be a recording of speech). In the video example provided, sustained harmonic tones appear to represent the movement of the *Mouth Piece* device as it approaches the camera, while the bell-like tones form isomorphic relationships with the moving text characters that make up the title 'Mouth Piece'.

Movie example 7 (experimental material from the writer's 2009 work *Abide With Me*) concerns the use of abstract video and referential audio displaying isomorphic relationships. In the example provided, not all the isomorphic relationships are intact (the example is flawed). The referential components of the audio and video materials fail to evoke a common schema – and thereby become concomitant. Interestingly, this does not prevent the intact abstract (parametric) isomorphic relationships from acting as catalysts in the process of integration. The result is a paradox – the production of material that appears to be both heterogeneous and homogeneous.

Movie example 8 (further experimental material from *Abide With Me*) concerns the use of abstract video and abstract audio displaying isomorphic relationships. A feature of this category is the focus on alignment of abstract video and abstract audio parameters. In the example provided the video materials (the four dots) were mapped in terms of brightness, onscreen height, and left/right (L/R) position. The data was then used to transform the accompanying audio materials (triangle waves) in terms of amplitude, frequency and panning. The result is the production of homogenous audiovisual materials.

To summarise at this point, materials that exhibit isomorphic relationships activate solitary schemas. Both referential and abstract qualities operate as catalysts in the process of integration. When isomorphic relationships fail, the resulting materials can exhibit both heterogeneous and homogeneous qualities. Furthermore, abstract isomorphic relationships may be created and controlled through the alignment of audio and video parameters.

Empirical studies have established that 'integrating visual and auditory stimuli serves the purpose of enhancing perceptual clarity' (Talsma et al. 2007: 679). Practice-led research concerning 'perceptual enhancement' has also been carried out in a number of high-information-flow situations, and experimental technologies have been trialled in various commercial settings such as the New York Stock Exchange, and in the visors of fighter pilots' helmets. At the heart of this research is the existence of two 'super-additive effects' that are known to occur where strong isomorphic relationships are observed. The first relates to holistic stimuli, the second concerns integration time – as Talsma explains: 'First, when the auditory, visual and audiovisual objects were attended, the P50 to the audiovisual stimuli was larger than the sum of the P50 activity for the auditory and visual stimuli…' and 'Their results, just as ours suggest that the multisensory integration effect only occurs early in time when both visual and auditory stimulus features can be constructed into a single coherent audiovisual object' (Talsma et al. 2007: 686, 688). Perhaps the momentary acuteness of sensation we experience when attending materials that exhibit strong isomorphic relationships can be attributed to these super-additive effects.

7. BUILDING ISOMORPHIC RELATIONSHIPS

An established method of building isomorphic relationships between abstract video and abstract audio materials is to:

- begin with the video footage;
- select the accompanying audio through a process of trial and error;
- time align the selected sounds with the video textures using onsets and durations; and
- edit the length of sounds to best suit the durations of the video textures.

Movie example 9 (an extract from an early version of *Abide With Me*) illustrates such an approach. The overall result is that isomorphic relationships tend to be formed for brief periods, fail at various junctures, become concomitant, and re-form at ensuing points of congruence. This is evident in the example provided, and is, in the view of the writer, an undesirable and unmanageable result.

In late 2008 a study concerning parametric audiovisual mapping strategies was begun by the writer. The hypothesis was that a 'system' (using a minimum of three media pairs) must be created to provide the illusion of a plausible isomorphic relationship. The practice-led study was spearheaded by the production concerns of the writer's 2009 work *Abide With Me*. The audiovisual parametric mapping strategy shown in Table 1 is drawn from the 10:04:09 entry of the process diary for that work (the items listed under the audio and video categories relate to modular objects used in Max/MSP).

A number of new research questions were raised in the process of investigating each of the nine fields identified. By way of illustration, in one experiment, snapshots of the five most active audio partials were mapped to colour. Various parametric assignments were arbitrary, while others were based on indiscrete experiential frameworks. For example, bright and dull sounds were intuitively matched to their counterpart colours, and colour-changing parameters were interpolated as partials changed. In relation to this mapping strategy, Cook states that 'we ourselves decide which colours and sounds will best serve the given assignment or emotion as we need them … [but also that] … a bright colour may correspond to a bright sound in a absolute sense' (Cook 1998: 52, 76). Although satisfactory (convincing) 'experiential' mapping solutions were obtained, further research in each of the nine fields identified would almost certainly lead to greater gains.

Random assignment and inversion of experiential values were two other methods investigated. The work of other researchers, such as Francisco Kröpfl, was also considered. He reports (in relation to his experiments) that: (A1) Dynamic correlation: envelopes related to displacement on depth and width axis of objects. (A2) Degrees of brightness related with pitch registers and spectral types. (B) Correspondence between structural properties of both dimensions (image and sound) such as: Regularity – irregularity, Homogeneity – heterogeneity, Continuity – discontinuity, Etc' (Kröpfl 2007: 90).

The initial outcome of the study was the development of three methods of parametric mapping – each featuring at least three of the nine spatial fields identified. The first two methods, 'audio-to-video' and 'video-to-audio', rely on measurement and reassignment of spatial parameters. The third, 'gestural control', relies on the real-time interpretation and 'performance' of spatial parameters. A brief summary of each technique follows (drawn from the 13:04:09 entry of the writer's process diary):

Video-to-Audio
(1) Begin with developed video.
(2) Consider video materials as numerous objects.
(3) Choose static audio materials to suit each object[i].
(4) Synchronise audio materials with video materials[ii] and bounce to file.
(5) Prepare (single object) video files.
(6) Track changes in video parameters and manipulate associated audio files (bounces) with visual data[iii].
(7) Assemble audio and re-synchronise with primary video source.

Audio-to-Video
(1) Begin with developed audio.
(2) Break audio materials into objects.
(3) Choose static video materials per object[i].
(4) Manipulate video with audio data[iii].

Table 1. An audio/visual parametric mapping strategy (based on experiential framework).

Common spatial parameter	Audio	Video	Comment
Duration/intensity (envelope)	Amplitude (L/R level maximum)	Frame number	i
X movement	L/R level difference	Matrix X position	ii
Y movement	Pitch (fundamental frequency)	Matrix Y position	iii
Z movement	High freq content (sum freq top 5 partials)	Scale frame 0–end	iv
Size	Centroid (average spectrum)	Scale matrix	
Shape	Spectral signature	Matrix positioning	
Elasticity of shape	Spectral signature over time (Δ top 4 p freq)	Select or generate	v
Material Quality 1	Spectral signature (ratio noise to tone)	Texture (plur/wake)	vi
Material Quality 2	Spectral signature (snapshot top 5 partials)	Colour (hsl2rgb)	
Material Quality 3	Spectral signature (high freq noise)	Solidness (sprinkle/wake)	vii

i. With regard to A-V techniques: using individual frames sorted from black to full luminance (e.g. 25 frames) as an alternative to using still images provides a means of introducing small variables (nuances) into video materials.
ii. With regard to A-V techniques: this parameter requires synthesising for mono sound sources – an effective strategy for stereo sounds can be to use sum and difference data (first or second order) generated by the addition/subtraction of left and right channels. This strategy is also valuable where the gestural qualities of sounds (spectral qualities) are used as a means of spatialising materials – of particular importance in multichannel electroacoustic music.
iii. This parameter is mapped according to the universal perception that high/low sounds come from high/low locations. Inversions can also yield interesting results.
iv. Based on close and distant variants of the same sound/image.
v. Here, the two concepts of 'spatial image' (abstract) and 'schema' (referential) are useful in assigning appropriate materials. Contemplation of the shape and size of resonant objects that may have 'created' the sound is also something to consider. To date, no effective 'generative' mapping strategy with regard to shape has been identified. At this time, the subjective (selection) approach remains the most satisfactory.
vi. The material quality is related to the rough/smooth spectral signature. This could be determined through analysis of partials.
vii. The solidness of the materials seems to be related to the abstract/referential continuum. Sounds are generally 'more abstract' than visual materials, or, to put it another way, visual materials are less ambiguous. The essence of this mapping problem seems to be that the referential/abstract continuums of sound and vision are scaled differently – although abstract visual images are abstract, they are generally more referential than their sonic counterparts. One means of abstracting these objects further might be to obscure them.

(5) Assemble video and re-synchronise with developed audio.
   (i) Select or record materials with the knowledge that certain parameters will not remain static after processing.
   (ii) Synchronised spectral changes in source 'recordings' are most desirable (preferable to spectral changes made during postproduction). Preview audiovisual materials in PT, and bounce audio materials per object or object component to file.
   (iii) This can be achieved in Max/MSP.

Gestural Control
(1) Contemplate the audio and video parameters that may benefit from either synthesis or adaptation using gestural control techniques.
(2) Custom design a gestural control interface (instrument) using interactive technologies such as Wii controllers, iCubeX sensors and/or video cameras (for video tracking). Develop the instrument for the sole purpose of effectively controlling the chosen parameter.
(3) Attend the associated audio or video media while 'performing' the counterpart in real time.
(4) Repeat steps 1–3 until at least 3 simultaneous media pairs have been produced. Alternatively, instruments may be designed to control more that one parameter simultaneously.

Movie example 10 (an extract of *Abide With Me*) illustrates the use of all three isomorphic media pairing strategies in sequence: gestural control, video-to-audio, audio-to-video, and gestural control. The principal 'dot' is 'morphed' between the three techniques to take advantage of the different qualities offered.

Two methods of gestural control will now be explained. In 'Dance', isomorphic relationships can be created by attending audio materials and 'acting out' visual counterparts. The impression is often one of reversal – that the movements of the dancer are controlling the audio. Leigh Landy takes advantage of this illusion in his work *I Conduct Electricity* (Landy 2000) by creating the impression that he is 'controlling' a series of pre-recorded sound materials (being played through stereo monitors) in various complex ways through the movements of his conductor's baton.

This principle (of reversal) is adopted in Movie example 10. An instrument was developed to allow manipulation of video footage in real time through the movements of a Wii controller. The associated audio was then attended, and the counterpart video was 'acted out'. The result is not one of subservience, but of congruence. This technique and others like it has many advantages over the analysis–reassignment-based parametric mapping strategies previously discussed – not the least of which is the ability to form complex systemic relationships in real time through the reassignment of parametric control data produced by intuitive movement.

A related method of gestural control is the real time synthesis of audio and video materials from a single data source. The writer's 2009 work *Eyepiece* makes use of an electronic device to track the unique eye movements of individual users and translate the data in real time into sounds and moving images. Mapping strategies similar to those listed in Table 1 are used throughout the work, and acute isomorphic relationships are realised as a result of the real-time interactivity. Movie example 11 demonstrates the *Eyepiece* device in operation. During the process of developing the work, an interesting discovery was made – that the nature of individual mapping strategies appears to be less important than the nature of the systems they constitute. Surprisingly, even inverted L/R assignments remain relatively plausible in the company of other parameters displaying dynamic isomorphic relationships. Movie examples 12–15 illustrate this phenomenon. The significance of the discovery is that the experiential framework (on which plausibility is based) appears to be far more adaptable than first anticipated. It may even be possible to evoke convincing isomorphic relationships through arbitrary or random assignments of three or more simultaneous parameters. Further investigation is required.

In summary, while video-to-audio and audio-to-video strategies tend to offer a level of precision barely obtainable elsewhere, strategies of gestural control tend to offer the most flexible, expressive and 'fun' solutions. Real-time 'instruments' may be designed and built to adapt, transform or synthesise materials at any phase of production in any given work. Furthermore, it appears that 'systems' (as opposed to rigid experiential frameworks) comprised of three or more dynamically assigned audiovisual parameters provide an effective means of maintaining convincing isomorphic relationships.

8. TRANSITIONS BETWEEN HETEROGENEOUS AND HOMOGENEOUS PLANES

Heterogeneity is experienced in its most acute form where two (or more) independent audio and visual features demand simultaneous attention (divided attention). Further down the scale, yet still within the realm of heterogeneous experience, is our voluntary capacity to selectively attend either audio or visual features within complex audiovisual environments. This ability is known as the 'cocktail party effect'. Further down still is our ability to perceive, yet ignore, major heterogeneous features as part of the involuntary process of integration (towards the homogeneous plane). Approaching the realm of homogeneity, we find materials that contain both heterogeneous and homogeneous features, and finally, at the extreme end of the homogeneous experience, we find perfectly integrated materials that induce only solitary schemas. The topic of this section concerns the middle ground – the agents that aid and disrupt the involuntary process of integration towards the homogeneous plane. Four contributors have been identified.

The first concerns the ability of the listener/viewer to apprehend the effects of the eight different types of audiovisual materials identified. As previously stated, media pairs tend to exhibit instantaneous combinations of concomitant and isomorphic relationships, and exactly how the combinations affect the process of integration varies greatly amongst listeners/viewers. Where one individual might perceive an isomorphic relationship that acts as a catalyst in the process of integration, another might experience the same materials in combination with a memory residue (perhaps from the programme note) that for him or her evokes a concomitant relationship – resulting in the masking of all other features. Such polar experiences are commonplace. Location, setting, and time of day are also powerful contributors.

The second concerns the effects of masking. Where tentative isomorphic relationships exist in the presence of other media pairs, the threat of masking is ever present. In these settings it is not the absence of isomorphic relationships, but our inability to perceive them that transforms our experience from the homogeneous to the heterogeneous. When masking effects are encountered, we tend to make several attempts to re-integrate the audiovisual materials for the purpose of 're-starting' the homogeneous experience. For the most part these attempts are unsuccessful.

The third agent identified is 'bandwidth' or 'bottlenecking'. Mari Riess Jones and William Yee elaborate: 'Many approaches to attention build upon an information processing metaphor. An individual is viewed as a communication system through which information flows. Attention is conventionally conceived either in terms of a bottleneck which occurs at certain stages in the flow, or in terms of certain processing limits…' (Riess Jones and Yee 1993: 72). In Movie example 10, 88 'dots' displaying isomorphic relationships are presented simultaneously – beginning with 1, and logarithmically increasing to 88 over a period of approximately two minutes. For the majority of listeners/viewers, only three or four dots can be attended at one time. At this juncture their experience is transformed from the homogeneous to the heterogeneous. Homogeneity is eventually restored (for one dot only) through the technique of 'amplifying' congruent audiovisual parameters.

The final point concerns the effects of interactivity. It appears that the primary and secondary feedback loops associated with interactivity help to maintain isomorphic relationships. In *Eyepiece* materials are initially perceived as acutely homogeneous. In playing back the recorded materials, the homogeneity is less pronounced, and finally, in the presence of other media pairs, the same materials begin to exhibit heterogeneous qualities. It is the action/inaction, and the location of the participant(s), that causes the transformation. Movie example 16 (an extract from a live interaction with *Eyepiece*) illustrates the effects of all four agents (to varying degrees).

9. SUMMARY AND CONCLUSIONS

I return to the paradox presented at the beginning of the article – that audiovisual materials, on one hand, offer the possibility of complementing one another, and, on the other, carry the risk of detracting from one another. It has now been determined that the conflicting statements are simple descriptions of homogeneity and heterogeneity. The paradox has been transcended.

In summary, the study has put forward a system for the comprehensive classification of media pairs (the cube). Eight examples have been provided to illustrate the functional qualities of each category (one from each corner of the cube). In attending audiovisual materials our natural tendency is not to divide our attention, but to integrate even the most heterogeneous of textures. Integration occurs when congruous relationships are perceived to exist between audio and video materials. Two types of relationships have been identified: 'concomitant' and 'isomorphic'. Concomitant relationships rely on the highlighting and masking that occurs when two schemas are simultaneously activated (overlaid), while isomorphic relationships act as catalysts in the process of integration by inducing the activation of solitary schemas. Detailed methods of creating both concomitant and isomorphic relationships have been proposed. These include: Aristotle's formulae, schematic extension, analysis-reassignment-based mapping strategies, and gestural-control-based mapping strategies. Finally, four agents that aid and disrupt the involuntary process of integration towards the homogeneous plane have been identified. These include: the perception of the listener/viewer, and the effects of masking, bandwidth (or bottlenecking) and interactivity.

The narrow findings of the study offer a model and vocabulary for considering the functionality of audiovisual materials. Although there are a number of other established constructs and borrowed terms available for discussing electroacoustic music with moving images (for example, those found in acousmatic electroacoustic music and filmmaking), there are few that address the issue of 'language' so directly. As such, the model for classifying media pairs (figure 1) may prove to be a valuable tool for analysis and teaching purposes. The wider findings of the study impact directly on the associated domains of live electronic music, and hyper-instrument design. The overarching discoveries include the principle of building functional relationships between audio and visual textures, and the practical means of gaining creative control over materials through parametric mapping solutions. Of particular interest is the ability to embed such transformational strategies in real-time performance-based scenarios where both audio and visual materials constitute the musical experience.

REFERENCES

Chion, M. 1994. *Audio-Vision: Sound On Screen*, ed. Claudia Gorbman. New York: Columbia University Press.

Cook, N. 1998. *Analysing Musical Multimedia*. New York: Oxford University Press.

Coulter, J. 2007. The Language of Electroacoustic Music with Moving Images. Proceedings of the Electroacoustic Music Studies Network International Conference Series: EMS07 12–15 June 2007, De Montfort University, Leicester, UK. http://www.ems-network.org.

Emmerson, S. 2007. Seeing (or not Seeing) the Loudspeaker; Seeing (or not Seeing) the Music. In F. Barrière and C. Clozier (eds.) *Relationships Between Audition and Vision in the Creation in Electroacoustic Music*. Academie Internationale de Musique Electroacoustique/Bourges, Institut International de Musique Electroacoustique de Bourges/IMEB, Bourges cedex, France. Volume VIII (2004–05), 85–8.

Kröpfl, F. 2007. Integrating Sound and Visual Image as Artform. In *Relationships Between Audition and Vision in the Creation in Electroacoustic Music*. Academie Internationale de Musique Electroacoustique/Bourges, Institut International de Musique Electroacoustique de Bourges/IMEB, Bourges cedex, France. Volume VIII (2004–05), 89–90.

Landy, L. 2000. I Conduct Electricity. Video file. In L. Landy and E. Jamieson (eds.) *Devising Dance and Music: I.D.Φ.X Idée Fixe Experimental Sound and Movement Theatre*. Sunderland: University of Sunderland Press. http://www.mti.dmu.ac.uk/ ~ llandy/dance.html, Intermezzo: key company ingredients (web link), Part2.mov, 17:30–19:50.

Landy, L. 2007. *Understanding the Art of Sound Organization*. Cambridge, MA: The MIT Press.

Mathews, M. 2007. Music and Video Discussion in Relationships Between Audition and Vision in the Creation. In *Relationships Between Audition and Vision in the Creation in Electroacoustic Music*. Academie Internationale de Musique Electroacoustique/Bourges, Institut International de Musique Electroacoustique de Bourges/IMEB, Bourges cedex, France. Volume VIII (2004–05), 94–5.

OED 2009. *The Oxford English Dictionary Online*, 3rd edn (J. Simpson, Chief Editor). Retrieved December 2009 from http://dictionary.oed.com.

Riess Jones, M., and Yee, W. 1993. Attending to Auditory Events: The Role of Temporal Organization. In Stephen McAdams (ed.) *Thinking in Sound: The Cognitive Psychology of Human Audition*. Oxford: Oxford University Press, pp. 69–112.

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

Talsma, D., Doty, T., and Waldorff, M. 2007. Selective Attention and Audiovisual Integration: Is Attending to Both Modalities a Prerequisite for Early Integration? *Cerebral Cortex* **17**(3): 679–90.