Audiovisual Composition

The relationship between sonic and visual material is complex. Essentially, both sonic and visual material yield effects of their own when experienced in isolation. These effects are difficult to discuss in themselves, let alone in combination. Chion states that when combined, new effects are apparent. This makes the process even more complex. Despite this, difficulties regarding the interpretation of audio and visual material do not prevent artists from exploring that material. However, audiovisual composition which exploits structural relationships by its nature rises out of a desire to understand the combined audiovisual effect. This article offers an argument for promoting the interdisciplinary study of audiovisual composition as a metadiscipline in itself that is neither Art/Film studies nor Music/Sonic arts.

In 'The Perception of Audio-Visual Composites: Accent Structure Alignment of Simple Stimuli' (Lipscomb, 2004), Lipscomb's results suggest a correlation between the synchronisation of audiovisual composites and the perceived effectiveness of the material. Where audio pulses correlate strongly with visual events, audiences (organised in ability groups according to background and training) sense a greater effectiveness[1].

> Calculation... revealed that subject ratings of synchronization and effectiveness shared a strong positive relationship (r = .96). Therefore, AV combinations that were rated high in synchronization also tended to be rated high on effectiveness and vice versa. (Lipscomb, 2005 p60)

This research shows that audiences perceive closely synchronised material as being more effective. However, Audiovisual works which exhibit strong structural links in this way are sometimes referred to as being guilty of 'Mickey Mousing' – the exact, and by implication, simplistic synchronisation of visual and sonic events. In his paper 'Insects, Urine and Flatulance: On the Radical Potential of Mickey Mousing' (Birtwistle, 2002), Birtwistle mounts a robust response to this criticism:

> The close matching of musical sound and image is seen in negative terms:
> "… because of the implication that exact illustration is a rather tedious and silly way to relate music and image." (Curtis, S in Altman, R 1992, p201) Mickey Mousing is poor practice. It is considered unsubtle, unnecessary and creates humour when none is required...(however) Mickey Mousing punctures the bubble in which western music has placed itself, forcing an acknowledgement of an 'outside', an other: in this case, the visual. Not only does Mickey Mousing destroy the notion of an isolated specificity, of an abstraction from all else, but it also introduces ideas of other kinds of structuration, other ways of considering structure, other ways of thinking music, and other ways of thinking about music. (Birtwistle 2002 p26)

So, according to Birtwistle, there is a desire to maintain a separation between musical and visual art which rises out of existing traditions. Criticism of closely synchronised material emanates from this perspective. In *Analysing Musical Multimedia* (Cook, 1998), Nicholas Cook argues that multimedia is ‘predicated by difference’ (Cook, 1998 p56), and the ‘duplication of information across sensory modes’ (Cook, 1998 p41) cannot be described as multimedia. If one is willing to take seriously the results of Lipscomb's experiments, it seems that it may be unwise to dismiss the effectiveness of closely synchronised material. Audiovisual composition, in fact, may rely on an understanding of this 'effectiveness' and the complexity of its operation. Perhaps Cook is right, and it is not evidence of multimedia.

1Defining exactly what is meant by 'effective' in this context is problematic, although for the purpose of this thesis it will be assumed that it equates to the memorability and/or subjective preference of an audiovisual event when compared to other less accurately synchronised events, as this appears to be Lipscomb’s distinction.

However, this does not mean that it is unsophisticated or lacking in value.

In the conclusion to his paper on Audiovisual relations, Lipscomb makes the point that although there is a proliferation of abundant audiovisual material in our everyday lives, very little research is being carried out to analyse the dynamics of the material.

> ..given the sociological significance of the cinematic experience, it is quite surprising that there is still only a small amount of research literature available addressing issues involved in the cognitive processing of ecologically valid audio-visual stimuli. (Lipscomb 2005 p65)

Recent research at the Shimojo Psychophysics Laboratory underlines the need for further research regarding audiovisual relationships. In 'Visual Illusion Induced by Sound' (Kamitani, Y, Shimojo, S, 2002), proof that audiovisual material is processed in combination, and that this combination alters the perception of the material with definite effects has been confirmed by MRI scans. In one experiment, a subject sees a black dot appear on a screen for one frame. When accompanied by two small blips, the subject sees two dots, one after the other, even though there is only one. This suggests that strongly synchronised material is effective in producing a type of experience which is distinct from the experience of images or the experience of sounds in isolation from one another. As such, it provides evidence for Chion's notion of added value, whilst underlining the fact that the structural relationship between sound and visual material is at the heart of audiovisual composition. In this way, the pejorative term 'Mickey Mousing' lacks authority with respect to the discussion of audiovisual composition, and although the process of composition may shift to and from heavily synchronised material, the effectiveness of synchronisation could be, as Birtwistle argues, truly radical.

Added value is a technique of combining sounds and images in order to generate a third *audiovisual* form, which modern corporations, advertisers and media companies utilise heavily in order to create aesthetic *identities*. The combination of time related, or even heavily synchronised abstract/graphic (and in many cases, synthetic) material with sonic material (including music) is a common technique for creating audiovisual iconography. Film distributors, TV companies, and other media organisations have audiovisual logos which help propagate their chosen image. The power of audiovisual relationships is being overlooked within the realm of research, while Audiovisual relationships remain at the heart of much modern communication and commerce.

> "...if some multisensory cell responds to a light flash in the upper right portion of the visual field, that cell will respond to a sound only if it too comes from the same vicinity. Additionally, when visual and auditory inputs occur simultaneously, a multisensory cell responds more strongly than when either input occurs alone." (Sekuler, R. & Blake, R. 1985, p104).

Research demonstrates that multisensory cells in the brain respond directly to audiovisual experience, as opposed to simply audio or visual experience, and that these cells respond more strongly when events occur simultaneously and/or appear to come from the same source. This gives further support to Chion's ideas, as well as the research carried out by Shimojo, Boltz et al. This helps to reinforce the case for the promotion of the metadiscipline of audiovisual composition.

Mick Grierson

London 2007