Audiovisual Coalescence: The Effect of Synchrony in Film Score

Andrew Rogers and Ian Gibson,

University of Huddersfield
andrew.rogers@hud.ac.uk

Abstract. The emotional response to synchronous instances between musical and visual elements within the cinematic experience awaits thorough empirical evaluation. Film composition is defined here as a genre of stereotypes, whose methodologies are not feasibly subject to redevelopment. As consequence, the research focuses on improving the enculturated functions of film music. Hypothesising that an increased frequency of synchronisation between musical and visual accents will generate a positive evaluative response from an audioviewer, the experimental results determine significant preference for frequent synchronisation. Viewer expectancy, multisensory integration and attention are used in conclusion of the findings. Argument is established for the necessary quantification of audiovisual accent saliency in refinement of audiovisual design.

Keywords: Emotion, Audiovisual, Music, Film, Television, Composition, Synchrony, Perception, Cognition, Auditory, Visual, Expectancy, Salience, Attention, Multisensory Integration

1 Enculturation

Upon listening to music, activations of the inferior frontal gyrus of the brain[1] evoke emotions [1]. The assumption that such perceptual influence is a fundamental export of music holds prominence in academia[2], yet this attaches emotion to an entity devoid of any such property [3]. However, such association is sufficiently prevalent that it is not limited to human or even human infants [4]. Cross species studies have shown that pitch and timbre can elicit consistent emotional responses [5].

Auditory information (and therefore music) gains this emotional potential via our perceptual mechanism of metaphor [6]. To avoid this process and listen to any sound as object (opposed to the indicative and communicative modes of listening which correlate to emotion [7]) requires significant training and effort [8]. The representational process of music as metaphor occurs in our perception and

---

[1] The Inferior Brodmann's area (BA) 44, BA 45, and BA 46. An area associated with music-syntactic analysis and working memory operations [1].

[2] Although this is an infrequently challenged suggestion, notable exceptions include Stravinsky and Hindemith. For further discussion see [2].

imagination [9], which is determined by our individual experiences and *enculturation* [10].

When music is utilised within the cinematic experience, stimulation and therefore perception is multisensory to the audioviewer. The consequence of which is significant, as the individual sensory modalities have sufficient influence to alter perception in the other [11] [12]. Music therefore amasses new metaphorical, perceptual and emotional potential within a multisensory environment such as film.

The Hollywood movie system's command of Western cinema has enculturated the target audience in their response to this amassed potential. Such dominance of a branch of popular culture from near conception has inevitably led to a cycle of repetition, resulting in prosaic works targeted at the mass market. Music composed for film has no immunity to this stalemate of proven formula. Cliché is commonplace and even the relatively small group of composers that dominate the mainstream industry are recognising the stagnation of creative development [13][3]. Association is the cause; conventionalising musical inference has solidified the methodologies of film scoring [15], generating a creativity paradox for composers. Resultantly academics often cite film composition as *ipso facto*, that is, inferior to concert music. This is to judge the work outside of its placement and purpose though. Music within film does not rely exclusively on purely musical signification. Cultural musical codes (incepted through enculturation) and cinematic codes determine an audioviewer's response [15] [16]. The banal design builds the semiotic complex of the film's overall purpose.

The functions of music within film have become determinable in their purpose by their audience. But accepting the functionality of established methodologies is to ignore the potential of their refinement.

2 Synchrony

The perceptual and emotional response to synchronous instances between musical and visual elements within the cinematic experience awaits thorough empirical evaluation. Research into multi-sensory integration demonstrates that one perceptual modality has the ability to enhance or bias another under particular temporal and spatial constraints [17]. Musical accompaniment to film is bound within such constraints.

"The extent to which sound activates an image depends on how it introduces points of synchronisation – predictably or not, variously or monotonously." Chion [8].

Academics have begun to direct their attention to the potentials of the highly specific audiovisual effects of film. Yet establishing viable methodology to explicate this is proving problematic [18]. Research by Lipscomb and Kendall [19] has utilised Osgood, Suci and Tennenbaum's [20] division of the semantic differential scale to evaluate audiovisual perception. Although the *Potency* and *Activity* aspects of the

---

[3] However, such composers are hypocritical in this analysis. Alberge [13] details Hans Zimmer's (Academy Award winning composer) praise for John Williams (multi Academy Award winning composer), but Williams' success can be largely attributed to remaining "true to the conventions of the form" i.e. non-progressive, recognisable compositional functions [14].

tripartite scale are convincingly addressed, the *Evaluative* impact remains under explained. Thereupon this study aims to define the evaluative impacts of varying musical accompaniment within the cinematic experience.

3 Empiricism

The working hypothesis for this research was that an increased frequency of synchronous instances between visual and musical accents would improve evaluative preference of the audiovisual product.

Previous research has substituted randomised musical content over visual stimuli to modify points of synchrony for analysis (Chion's 'Forced Marriage' experiments for example [8]). However, with varying musical content the experimental procedure would lean towards a musical preference test, rather than an audiovisual one. To resolve this experimental problem the musical examples would be identical throughout[4]. Evidently this is impossible due to the alteration of musical structure necessary to investigate the variable, instances of synchronisation. Work by Lipscomb [19] has addressed this issue by time stretching compositions to alter their interaction with the visual. Altering the tempo through mechanical means though arguably has effect on the emotional impact (herein *pathos*) of the music, biasing subject evaluations. It also has the undesirable effect of utilising music intended for different visual cues than those composed for, undermining the work of the composer.

In this experiment a solution to the problem has been achieved by maintaining all emotional stimulants of the compositions used, whilst still introducing subtle variations to distribute points of synchronisation. Great care was taken in maintaining all musical elements across compositional variations[5] to maintain pathos.

3.1 Experimental Design

Nine audiovisual scenes were designed for evaluation by 48 subjects via a verbal scaling procedure. This was a free ranging scale measured to 101 points designed to avoid both range equalising and dumping bias. Ambiguous terminology denoting more to less suitable audiovisual cohesion defined the scale to sufficiently avoid anchoring terminology. Three visual scenes were utilized, each with three musical variations. The musical variations were minor alterations of a composer's initial response to the scene to allow for highly, partially, and nonsynchronous audiovisual interaction whilst maintaining equal pathos across visual scene repetitions.

The three musical variations for each of the three visual scenes were composed to comply with each of the following:

---

[4] Or entirely randomised to omit musical genre influence and focus on the audiovisual interaction. This would require an extremely broad data set.

[5] Including harmony, melody, timbre, tempo, instrumentation, mix balance, perceived loudness and so forth.

- Nonsynchronous. No constructed points of synchronisation between visual accents and musical accents[6].
- Highly Synchronous. Many constructed points of synchronisation between visual accents and musical accents.
- Partially synchronous. Some constructed points of synchronisation between visual accents and musical accents.

Initially, one of the three methods mentioned above was employed to compose musical accompaniment to the video clip. Of the three clips requiring musical variation under assessment each was primarily composed using one of the three methods. Each of the methods was used once for primary composition to avoid bias from this potentially advantageous compositional factor. Once a composition was established, a rework into the two remaining forms was undertaken. This would therefore involve manipulation of timing to alter points of synchronisation from the composition.

3.2 Results

The partially and fully synchronous audiovisual scenes were preferable over their nonsynchronous counterparts in all examples. Clips two and three (Fig. 2 and Fig. 3) provided significant ($p < .05$) results, whereas clip one (Fig. 1) produced no statistically significant result. These results support those from other multisensory and audiovisual studies, but importantly establish them within the ecologically valid environment of 'real world' film and dedicated musical score.
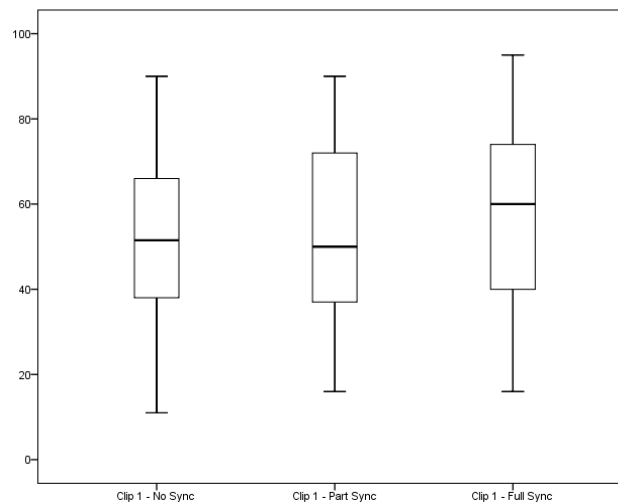

Fig. 1. Clip 1, subject evaluation across synchronous variations.

---

[6] Fortuitous instances of perceivable synchronisation were not actively avoided.
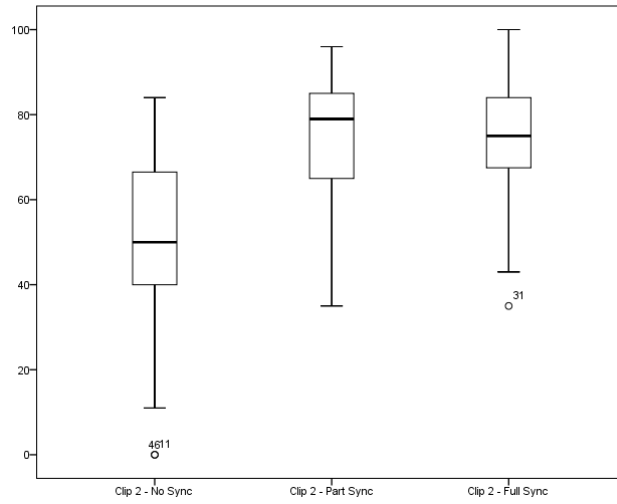
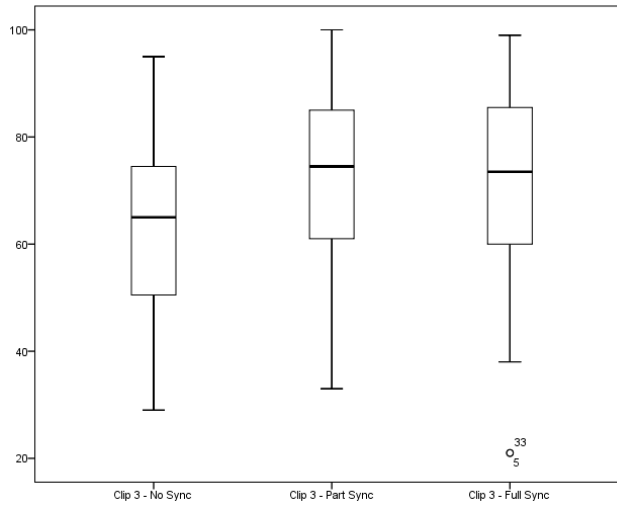**Fig. 2**. Clip 2, subject evaluation across synchronous variations.

Fig. 3. Clip 3, subject evaluation across synchronous variations.

4 Conclusion

The exploitation of our senses is the driver for our culture, entertainment and education. The degree of sensory invigoration and emotional impact heightens when stimuli are cumulative. This study has demonstrated an evaluative preference towards synchronous audiovisual material over non synchronous audiovisual material in a cinematic context. However, the frequency of synchrony did not directly correlate with subject evaluation. Therefore the hypothesis that the frequency of synchronous instances between visual and musical accents would *improve* evaluative preference of the audiovisual cannot be confirmed. The results do allow rejection of the null hypothesis though, confirming that the frequency of synchronous instances between visual and musical accents does have a significant impact on an audioviewer's evaluative preference.

Conclusion can be drawn initially through audioviewer *expectancy.* The required fluctuations in tempi and time signatures in the experimental design consequently input deviations from commonly utilised rhythmic structures (these being consistent and even rhythms to the Western enculturated subject group). Laurel [4] references musicologist Meyer [21] who established that the most emotionally charged aspects of music are associated with low-probability events[7]. As the evaluative response is a descendent of the emotional [23], and this response lay with the more synchronous material, it can be concluded that the preference was a measurement of emotional stimulation through a density of low-probability events. Huron's [24] theory that emotions are evoked from deviations in expectation due to the mind being "wired for expectation" as a key survival instinct further support this argument.

The above conclusion focuses on auditory functions affecting response, yet this outcome must be considered in a multisensory context as we have both auditory and visual elements. As Koelewijn et al. [17] explain, "When auditory and visual events are presented at roughly the same time and location they tend to integrate... This integration can lead to an increased saliency and can draw attention in cases in which individual stimuli would be less effective." *Attention* refers to processes that accommodate selective processing of incoming stimuli, where multisensory attention results in a concomitant shift of attention in the other senses at work. Thus, the unexpected musical structures can heighten attention both visually and aurally which again accounts for the higher emotional impact.

These conclusions do not account for the fact that statistically significant impact of synchronous compositional variation was measured in only two out of the three examples evaluated by subjects. This is an enlightening result though, as it allows for analysis of the differences between the significant and non-significant clips. Research indicates that it is not only the frequency, but also the *salience* of synchronous events that affects perception [17]. On analysis of the interaction of audiovisual components across the clips, this appears to be the deciding factor, as clip 1's (Fig. 1) synchrony salience was less than that of clip 2 and clip 3.

Quantifying the significance of synchrony salience becomes the next step in understanding the interaction. Deviations in the plethora of musical functions can be used to quantify a separation from the previous, thus creating a measurable musical

---

[7] This has been further validated by research using electroencephalography [22].

accent. Quantification of visual accents relies upon a visual equivalent of the stated. Further research is necessary to achieve quantification of these accents to equate a relationship that can be utilised in the composition of appropriate audiovisual synchronicity.

References

1. Koelsch, S., et al., Investigating emotion with music: an fMRI study. Hum Brain Mapp, 2006. **27**(3): p. 239-50.
2. Mew, P., The expression of emotion in music. The British Journal of Aesthetics, 1985. **25**(1): p. 33-42.
3. Zangwill, N., Music, Metaphor, and Emotion. The Journal of Aesthetics and Art Criticism, 2007. **65**(4): p. 391-400.
4. Laurel, T., The emotional origins of music. Physics of Life Reviews, 2010. **7**(1): p. 44-45.
5. Morton, E.S., On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. American Naturalist, 1977. **111**: p. 329-333.
6. Aragno, A., Meaning's Vessel: A Metapsychological Understanding of Metaphor. Psychoanalytic Inquiry, 2009. **29**(1): p. 30-47.
7. Kane, B., L'Objet Sonore Maintenant: Pierre Schaeffer, sound objects and the phenomenological reduction. Organised Sound, 2007. **12**: p. 15-24.
8. Gorbman, C. and M. Chion, Audio-vision: sound on screen1994, New York: Columbia University Press.
9. Peacocke, C., The Perception of Music: Sources of Significance. The British Journal of Aesthetics, 2009. **49**(3): p. 257-275.
10. Morrison, S.J., S.M. Demorest, and L.A. Stambaugh, Enculturation Effects in Music Cognition: The Role of Age and Music Complexity. Journal of Research in Music Education, 2008. **56**(No. 2): p. 118-129.
11. Amedi, A., R. Malach, and A. Pascual-Leone, Negative BOLD Differentiates Visual Imagery and Perception. Neuron, 2005. **48**(5): p. 859-872.
12. Recanzone, G.H., Auditory influences on visual temporal rate perception. Journal of neurophysiology, 2003. **89**(2): p. 1078.
13. Alberge, D., Modern film scores are terrible, say composers, in The Times2008: London (UK), United Kingdom, London (UK). p. 31.
14. Scheurer, T.E., John Williams and film music since 1971. Popular Music and Society, 1997. **21**(1): p. 59-72.
15. Lipscomb, S.D.T., D. E. , ed. The role of music communication in cinema. Music communication, ed. R.M. D.Miell, & D. Hargreaves2005, Oxford University Press: Oxford, UK 283-404.
16. Gorbman, C. and I. British Film, Unheard melodies: narrative film music1987, Bloomington: Indiana University Press.
17. Koelewijn, T., A. Bronkhorst, and J. Theeuwes, Attention and the multiple stages of multisensory integration: A review of audiovisual studies. Acta Psychologica, 2010. **134**(3): p. 372-384.
18. Joost, G., S. Buchmuller, and R. Englert, Audio-visual Rhetoric: Visualizing the Pattern Language of Film, in Undisciplined! Design Research Society Conference 20082008: Sheffield Hallam University, Sheffield, UK.
19. Lipscomb, S.D. and R.A. Kendall, Perceptual Judgement of the Relationship between Musical and Visual Components in Film. Psychomusicology, 1996. **13**(Journal Article): p. 60-98.

20. Osgood, C.E., G.J. Suci, and P.H. Tannenbaum, The measurement of meaning1957: University of Illinois Press.
21. Meyer, L.B., Emotion and meaning in music1961, Chicago: University of Chicago Press.
22. Trainor, L.J. and R. Zatorre, The neurobiological basis of musical expectations: from probabilities to emotional meaning in Oxford handbook of music psychology S. Hallam, I. Cross, and M. Thaut, Editors. 2009, Oxford University Press: Oxford. p. 171-183.
23. Clore, G.L. and J.R. Huntsinger, How emotions inform judgment and regulate thought. Trends Cogn Sci, 2007. **11**(9): p. 393-9.
24. Huron, D., Sweet Anticipation : Music and the Psychology of Expectation2006: MIT Press.