Emotional Impact of Musical/Visual Synchrony Variation in Film

Andrew Rogers

University of Huddersfield, United Kingdom
andrew.rogers@hud.ac.uk

ABSTRACT

The emotional impact of synchronous musical and visual prominences 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 significant redevelopment. As consequence, the research focuses on improving components of the audience recognisable functions of film music. Subjects graded cinematic clips with musical elements that varied in their synchronous interaction with visual prominences. A positive response to more frequent synchronisation between music and film was concluded. Perceptual expectancy, attention and multisensory integration are principal in analysis of the findings.

I. Perception, Enculturation and Cliché

Upon listening to music, activations of the brain evoke emotions [1]. Conjecture that emotional influence is a fundamental export of music holds prominence in academia[2], yet this attaches emotion to an entity devoid of any such property (Zangwill, 2007). However, such attachment is sufficiently prevalent that it is not limited to human or even human infants (Laurel, 2010). Cross species studies have shown that pitch and timbre can elicit consistent emotional responses (Morton, 1977).

Aragano (2009) explains that auditory information (and therefore music) gains this emotional potential via our perceptual mechanism of metaphor. To avoid this process and listen to sounds as disassociated objects[3] requires significant training and effort (Gorbman & Chion, 1994). The overriding representational process of music as metaphor occurs in our perception and imagination (Peacocke, 2009), which is determined by our individual experiences and *enculturation*[4] (Morrison et. al, 2008).

Stimulation becomes multisensory within the cinematic experience as information is received through both visual and auditory sensory modalities. The consequence of this is significant, as the individual sensory modalities have sufficient influence to alter perception in the other (Amedi et al., 2005) (Recanzone, 2003). Music therefore amasses new metaphorical, perceptual and emotional potential within the multisensory environment.

The Hollywood movie system's command of western cinema has enculturated the target audience in their response to this amassed potential. Such consistency in response is reflected in the industries cycle of repetition, resulting in prosaic works targeted at the mass market[5]. Music composed for film has no immunity to the 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 creativity[6]. Association is the cause; enculturated musical inference has solidified the methodologies of film scoring (Lipscomb, 2005), generating a creativity paradox for composers. Resultantly academics and critics often cite film composition as *ipso facto*, that is, inferior to concert music (Wierzbiki, 2009). A comparison between the two distinct art forms is a redundant process though. Music within film does not rely exclusively on the musical signification that ties it to concert music or additionally popular music which has established itself within the film soundtrack[7]. Cultural musical codes[8] *and* cinematic codes (of which the interaction of music and image is component in establishing) determine a film viewer's response (Gorbman, 1987). The criticised banal musical design common in music composed for film builds the semiotic complex of the film's overall purpose through implementation of recognisable schema. But accepting the efficacy of established methodologies is to ignore the potential of their refinement.

II. 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 (within which musical accompaniment to film is bound) (Koelewijn, 2010).

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

---

1 Activations take place within the Inferior Brodmann's area, BA 44, BA 45, and BA 46. This is an area associated with music-syntactic analysis and working memory operations (Koelsche et al., 2006).

2 This is an infrequently challenged suggestion, yet Hindermith argues well in opposition. For further discussion see (Mew, 1985).

3 As opposed to the indicative and communicative modes of listening which correlate to emotion (Kane, 2007).

4 Enculturation in the context of the cinematic experience refers to an accumulated appreciation of the repeated conventions of film. In this project, examples from Western cinema were used as the source material on a Western enculturated subject group.

5 The top six grossing films worldwide in 2011 were sequels (four out of six in 2010). Repetition is highly successful.

6 However, such composers are hypocritical in this analysis. Alberge (2008) 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 (Scheurer, 1997).

7 The films of Quentin Tarantino expertly demonstrate the potential of popular music in film.

8 This includes any associations drawn from previous incarnations of the music if not composed specifically for the film.

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 (Joost et al., 2008). Research by Lipscomb and Kendall (1996) has utilised Osgood, Suci and Tennenbaum's (1957) division of the semantic differential scale to evaluate audiovisual perception. Although the *Potency* and *Activity* aspects of the tripartite scale are convincingly explored, the *Evaluative* impact remains under-explained. This study aims to explore the evaluative impacts of varying the *interaction* of musical and visual content within the cinematic experience.

III. Empericism

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). However, the significant diversity in the variable can result in the evaluation becoming more a musical preference test as opposed to an audiovisual one. Work by Lipscomb (1999) has considered this issue by time stretching compositions to alter their interaction with the visual. The problem with this methodology is that it belittles the intention of the film music composer, already shown by Lipscomb to be a key determining factor in assessment of preference. An alternative solution to the problem has been employed in this research to maintain both musical structure and composer intention across compositional variations whilst still introducing subtle variations to distribute points of synchronisation

A. Experimental Design

Three visual clips were utilized, selected for their diversity in cinematic style to effectively evaluate a potential overall trend[9]. Each clip had three musical composition variations. One of these was an original composition in reaction to the film clip. The remaining two variations were minor alterations of the composer's initial response to the clip[10] to distribute the synchronous interaction of music and visual elements. Points for synchronisation were determined from the visual by sudden visual edits (typically changes in camera angle) or salient temporal visual events in longer scenes (such as physical blows in a fight scene).

The knowledge that composer intended pairing of musical cues to their designated visual is deemed preferable by a majority of subjects creates a potential bias in compositional revisions. Therefore it was important to distribute the methodology of initial composition. This process is detailed in table 1.

Table. 1. Distribution of compositional process to avoid bias.

Composition	Clip 1	Clip 2	Clip 3
Initial	Most synchronous	Least synchronous	Partially synchronous
Revision A	Partially synchronous	Most synchronous	Least synchronous
Revision B	Least synchronous	Partially synchronous	Most synchronous

Subtle variations in tempi (to a maximum of +/- 4 beats per minute) were utilized to distribute synchronous instances between the musical and visual. Alterations in time signature and silence duration further aided the revision of synchronous distribution. All other musical functions and sonic production (such as melody, harmony, mix, perceived loudness and so forth) were maintained to the utmost possible extent to ensure compositional variations were as perceivably similar as possible.

Figure 1 details the process for the three compositional variations used for clip three in the experiment, showing the tempo track across a timeline (with track specific musical bar counter and time signature indicator) with arrows to represent the distribution of synchronous instances (downward arrows represent sync on a musical downbeat, upward and diagonal arrows represent up and offbeat sync respectively – blurred arrows represent loose/fortuitous sync[11]).

The total of nine audiovisual clips were audioviewed[12] by 48 subjects who gauged their opinion via a verbal scaling procedure realised on 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 (sufficiently avoiding anchoring terminology). The potential for interpretation of the *suitability of cohesion* as audiovisual synchrony by subjects was plausible, yet any mention of synchrony was avoided in the procedure. The experimental design ensured that even if subjects became aware that synchrony was the variable, they were still considering whether they found this preferential as an audiovisual presentation or not, maintaining the focus of the study[13]. The clips were shown in a randomised order with the inclusion of a repeated example to check for results validity within individual subjects.

[9] Of the three clips, one contained abstract visual imagery (50s duration), and the other two were narrative, one with dialogue (70s duration), and the other without dialogue (90s duration). The clips were from western cinema productions released in 1966, 1993 and 2003 to western audiences. The composed soundtracks lasted the duration of the clips.

[10] Compositions were produced by the author of this paper, Andrew Rogers.

[11] Sync was considered precise if within the range of one frame of video. Loose/Fortuitous sync labelled instances where music and video aligned within one to three frames.

[12] To "watch a film" is the common phrase regarding the cinematic experience. However, this omits the importance of sound, clarified here by the definition of 'audioview'. The audioviewing took place in a lecture theatre with a projector for the display and high end studio monitors for audio output.

[13] Results did not entirely favour the most synchronous examples, further eliminating the possibility of this misinterpretation of instruction.

B. Results

The partially and fully synchronous audiovisual clips were preferable over their nonsynchronous counterparts in all examples. Clips two and three (Fig. 3 and Fig. 4) provided significant ($p < .05$) results, whereas clip one (Fig. 2) produced no statistically significant result[14]. 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.

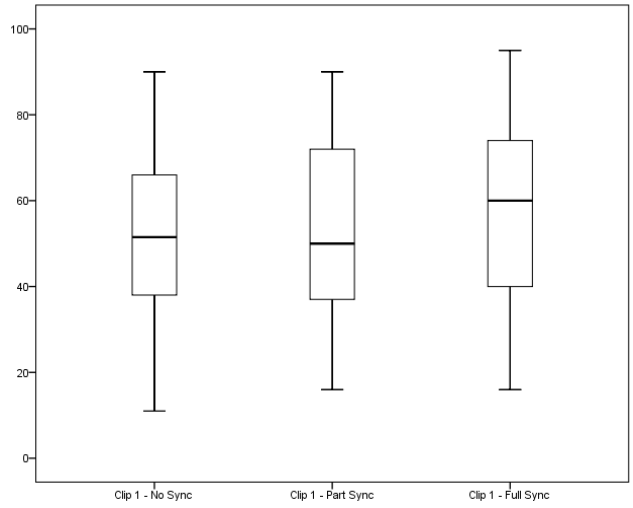

Fig. 2. Clip 1, subject evaluation across synchronous variations.

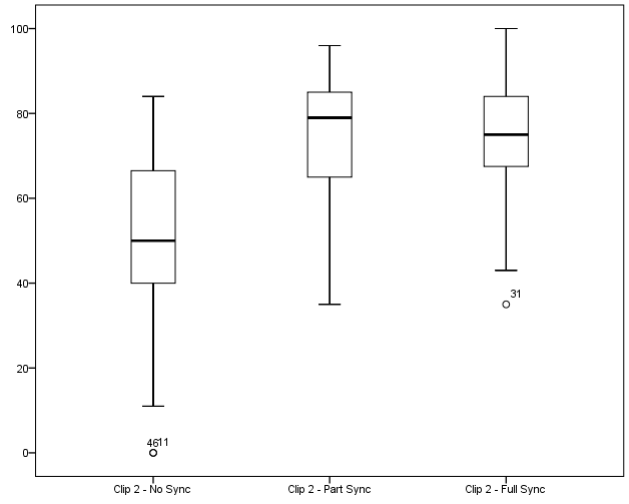

Fig. 3. Clip 2, subject evaluation across synchronous variations. Circles are outliers in the data.

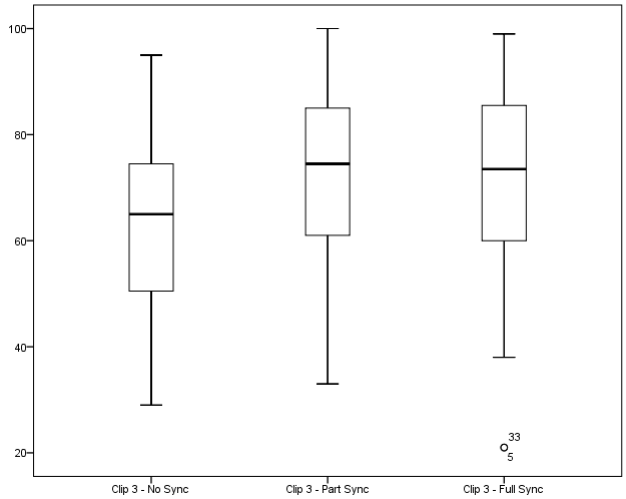

Fig. 4. Clip 3, subject evaluation across synchronous variations. Circles are outliers in the data.

IV. 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, confirming that the frequency of synchronous instances between visual and musical accents does have a significant impact on an audioviewer's evaluative preference in all instances tested.

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 references musicologist Meyer (1961) who established that the most emotionally charged aspects of music are associated with low-probability events[15]. As the evaluative response is a descendent of the emotional (Clore & Huntsinger, 2007), and the positive response lies with the more synchronous material, it is possible that the preference was a measurement of emotional stimulation through a density of low-probability events. Huron's (2006) documentation that emotions are evoked from deviations in expectation due to the mind being "wired for expectation" as a key survival instinct further supports this argument.

The above conclusion focuses on auditory functions affecting response, yet this outcome must be considered in a

[14] Although the result for clip was consonant with the other two clips, the results failed to reach the $p < .05$ level of significance.

[15] This has been further validated by research using electroencephalography (Trainor and Zatorre, 2009).

multisensory context. As Koelewijn et al. 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 the 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. Koelewijn et al. indicate that it is not only the frequency, but also the salience of synchronous events that affects perception. On analysis of the interaction of audiovisual components across the clips, this appears to be the deciding factor, as clip 1’s (Fig. 2) synchrony salience appears to be less than that of clip 2 and clip 3. Quantification of synchrony salience now becomes an essential factor in understanding the implications of musical to visual synchrony in film score.

REFERENCES

Alberge, D. (2008). Modern film scores are terrible, say composers, *The Times*, p. 31. Retrieved from http://search.proquest.com/docview/319845230?accountid=11526

Amedi, A., Malach, R., & Pascual-Leone, A. (2005). Negative BOLD Differentiates Visual Imagery and Perception. *Neuron, 48*(5), 859-872. doi: 10.1016/j.neuron.2005.10.032

Aragno, A. (2009). Meaning's Vessel: A Metapsychological Understanding of Metaphor. *Psychoanalytic Inquiry, 29*(1), 30-47. doi: 10.1080/07351690802247021

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends Cogn Sci, 11*(9), 393-399. doi: 10.1016/j.tics.2007.08.005

Gorbman, C., & British Film, I. (1987). *Unheard melodies: narrative film music*. Bloomington: Indiana University Press.

Gorbman, C., & Chion, M. (1994). *Audio-vision: sound on screen*. New York: Columbia University Press.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*: MIT Press.

Joost, G., Buchmuller, S., & Englert, R. (2008). *Audio-visual Rhetoric: Visualizing the Pattern Language of Film*. Paper presented at the Undisciplined! Design Research Society Conference 2008, Sheffield Hallam University, Sheffield, UK.

Kane, B. (2007). L’Objet Sonore Maintenant: Pierre Schaeffer, sound objects and the phenomenological reduction. *Organised Sound, 12*, 15-24. doi: 10.1017/S135577180700163X

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica, 134*(3), 372-384.

Koelsch, S., Fritz, T., DY, V. C., Muller, K., & Friederici, A. D. (2006). Investigating emotion with music: an fMRI study. [Research Support, Non-U.S. Gov't]. *Hum Brain Mapp, 27*(3), 239-250. doi: 10.1002/hbm.20180

Laurel, T. (2010). The emotional origins of music. *Physics of Life Reviews, 7*(1), 44-45. doi: 10.1016/j.plrev.2010.01.010

Lipscomb, S. D. (1999). Cross-modal integration: Synchronization of auditory and visual components in simple and complex media. *The Journal of the Acoustical Society of America, 105*(2), 1274.

Lipscomb, S. D. T., D. E. (Ed.). (2005). *The role of music communication in cinema.* Oxford, UK Oxford University Press.

Lipscomb, S. D., & Kendall, R. A. (1996). Perceptual Judgement of the Relationship between Musical and Visual Components in Film. *Psychomusicology, 13*(Journal Article), 60-98.

Mew, P. (1985). The Expression of Emotion in Music. *The British Journal of Aesthetics, 25*(1), 33-42. doi: 10.1093/bjaesthetics/25.1.33

Meyer, L. B. (1961). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Morrison, S. J., Demorest, S. M., & Stambaugh, L. A. (2008). Enculturation Effects in Music Cognition: The Role of Age and Music Complexity *Journal of Research in Music Education, 56*(No. 2), 118-129.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*: University of Illinois Press.

Peacocke, C. (2009). The Perception of Music: Sources of Significance. *The British Journal of Aesthetics, 49*(3), 257-275. doi: 10.1093/aesthj/ayp016

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of neurophysiology, 89*(2), 1078.

Scheurer, T. E. (1997). John Williams and film music since 1971. *Popular Music and Society, 21*(1), 59-72. doi: 10.1080/03007769708591655

Trainor, L. J., & Zatorre, R. (2009). The neurobiological basis of musical expectations: from probabilities to emotional meaning In S. Hallam, I. Cross & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 171-183). Oxford: Oxford University Press.

Wierzbicki, J. E. (2009). *Film music: a history*. New York: Routledge.

Zangwill, N. (2007). Music, Metaphor, and Emotion. *The Journal of Aesthetics and Art Criticism, 65*(4), 391-400.

.

Fig. 1. Points of synchronization across compositional variations for Clip 3