Roger Alsop and Paul Fletcher
Victorian College of the Arts
234 St Kilda Rd
Southbank, 3006
Australia
r.alsop@vca.unimelb.edu.au
p.fletcher@vca.unimelb.edu.au

Weaving Relationships between Sound and Image in Improvised Performance

Abstract

Integrating video and audio elements in improvised performance can be problematic in that the conventions of both media cause expectations in both the improvisers and their audience. These problems are magnified when more than one improviser is at play. A system that has roots in musical, narrative, and/or linguistic structures can be a starting point from which to form semantic zones. Within these zones, improvisers are free to react to the visual and aural elements resulting from the improvisation, while influencing those elements as the improvisation develops.

This paper discusses methods of building semantic zones; possible structural analogies between music, narrative, and language; and the development of an improvisational system based on those precepts.

Introduction

It can be argued that humans like to make sense of the events they witness or are involved in; that we look for a meaning or logic that fits our internal paradigms and hegemonies, and that we can understand and relate to in some way. This desire for understanding in music and film can be seen in the number of journal articles, where much of the discussion centres on attempts to better explain, and therefore understand, the affect and effect of these art forms. It is important to acknowledge that understanding is personal, and can be emotional, intellectual, spiritual, or any other kind.

When developing an understanding of an art work, particularly in temporally based forms such as music or film, the viewer may form a story or narrative on which they draw to develop an understanding and opinion of the artwork. When discussing the use of film as an educational tool Gary Berg saw that "Viewers of film clips tend to try and put a narrative structure onto the clips even if none is intended"(Berg 2000). It is possible that the same occurs when an audience is listening to music. This need for, and method of, understanding may be either inherited, trained, or, most likely, a mix of the two.

Margaret Boden describes a successful creative artwork as containing:

> "An arresting metaphor or poetic image, an unpredicted twist of the plot, a novel style of music, painting, or dance...all these unexpected things amaze and delight us."

she precedes this sentence with

> "Devotees of the humanities expect to be surprised."(Boden 1995)

Here Boden describes one of the paradigms through which ideal artists, and perhaps consequently an ideal arts audience, engage with an art work. What she infers is that there are paradigms that the viewer expects and through which they initially filter their understanding of the art work. These are confronted, subverted or indeed surprised by events that do not fit those paradigms.

David Bordwell and Kristin Thompson expand on this idea, saying:

> "Like emotion, meaning is important to our experience of artworks. As an active perceiver, the spectator is constantly testing the work [he or she is viewing] for larger significance, for what it says or suggests. The sorts of meanings that the spectator attributes to the work may vary considerably."(Bordwell 2001).

This constitutes a search for understanding beyond the work itself, fitting the work into a familiar story or methodology, and perhaps then seeing some relevancy to themselves.

Perhaps the artist engages and communicates with their audience by coaxing understanding from and within them. He or she may do this by defining the methodologies through which the audience members create their own communally held and idiosyncratic understandings. The diversity of this understanding may or may not bear any relation to the artist's intentions, and it is this diversity that often gives the work strength and relevance to different cultures at different times.

By fitting into conventional methods of understanding, the viewer may understand and assimilate the information without having to confront or reconstruct those methods. It is through the text that the work generates its affect. While the affect of music is not as easy to describe within textual paradigms, there are many conventions used by composers to guide an audience's understanding, and that audiences use to gain a shared understanding. These are usually in the non-textual/emotional realm, and are then described or conveyed in the textual/intellectual realm.

Animation is a field of the film-based arts that can fit well in both camps. Here it is possible to have images

that may be more easily understood through non-textual discourse. An example is the image Totoro, seen in the animation *My Neighbor Totoro (*Miyazaki 1988*)*, a furry bear like creature created by Hayao Miyazaki. As a character Totoro does not have much narrative function in the development of the plot other than to be there as the plot develops, forming a benign foil for other characters. A similar character is Boo Radley, the mostly silent character in *To Kill a Mockingbird* by Harper Lee; both function as informants to the plot without influencing it until the climax or resolution. Both characters could be seen to hold a similar semantic resonance, fitting into similar zones that he audience interprets and reinterprets as the work develops.

Semantic zones

Here we look at constructing semantic zones; areas of visual and sonic objects of similar semantic resonances that are defined in and by the work. These zones form pools of sonic and visual elements from which improvisers draw to reconstruct events, developing outcomes that redefine those elements in performance.

A semantic zone could be considered as a sector of influence and understandings within a kinetic, visual, textual or sonic language: or a set of conventions and expectations that are used to produce agreed or shared meanings and understandings.

This is equally the case in differing art forms, contexts and media. While the language may be specific and traceable to a particular artistic discipline and form, it can be seen to interact in varying ways and relationships with other language areas or zones. The relations between differing semantic zones may be parallel, entwined or overlaid.

Music is often described by its appeal to emotional and abstract meanings, and film with more concrete representations and information. Animation, however, is a form of film that can more easily fit within the same semantic spheres as dance and music, particularly in its more experimental formats.

The animated character Totoro is an example of a semantic zone; it is an image that creates a certain resonance, shared to a greater or lesser degree, depending on enculturation, by each viewer. Other, more pedestrian and less constructed images that hold shared semantic resonances may include doors, where the zone may be "transition"; lakes, which may imply "calm" or "depth"; telephones, "disembodied intercourse", knives, "danger"; trains, "journey" or "departure/arrival" and so on. The artist constructs the methodologies through which the viewer interprets these objects, thus creating semantic zones from which they build their own story.

Barthes discusses the role of a telephone in James Bond movies, describing it as "ringing during night duty at Secret Service headquarters – *Bond picked up one of the* ***four*** *receivers* [my bold] … indicating a concept necessary to the story (that of a highly developed bureaucratic technology) … [and demonstrating how word units such as *four* extend] beyond the level of denotation"(Barthes 1980). This is what the telephone does due to the number of them. This is an example of the artist, the film maker in this case, creating a context for an object that the viewers interpret into their own, often shared, understanding of the object and to a larger extent informing the entire work.

In music there are a number of semantic zones that include the harmonic, intervallic, timbre, rhythmic/tempo structures that elicit specific emotional responses. That these structures also elicit specific emotional responses makes the study of music, from the practical/theoretical to the ineffable, intriguing.

At a more idiomatic level the meanings that are given to triads and harmonic progressions within western musical traditions are an example of a different type of semantic zone. Here, from informal, colloquial points of view and conventions, the minor chord may be seen as representing sadness or similar emotions, and the major chord as representing happier emotions. These particular sounds, the minor and major chords, can have two semantic functions, the emotive one mentioned first and the second one that describes propulsion through cadential forms. In this case, the major chord can act as a point of stability of movement, depending on its context; the minor chord can have similar, though perhaps less potent, functions. This is an example of the one object fitting a number of semantic zones, with the understanding of the object being based on the context through which it is to be considered.

The liturgical Mass is an example of musical and textual semantic zones working in combination; here the narrative structure of the text is reflected in musical events; each of the six sections requiring music that reflects the semantic properties of the text. In this process the meanings of the musical objects heard are informed by the text, and the understanding of the text amplified by the music. This creates a loop for the audience, where the music is based on the text, which in turn creates an understanding of the music. As these elements coalesce they re-inform each other, creating a whole greater than the sum of its parts.

The film *Revolver* (Bergqvist, Martti Ekstrand et al. 1993) is another example of elements reconstituting and re-informing each other. Here a series of animated sequences or modules loop repeatedly. Its structure is akin to the musical form of theme and variations within the Minimalist style of repeating loops. The progression of a narrative does not appear to be a strong factor in *Revolver*. While the repetitions would intuitively produce a sense of stasis, the rhythm within each visual loop, the visual activity therein, and the timing of when new loops begin, provides a forward momentum.

Creating semantic zones

Bordwell and Thompson present four types of meaning, ranging from the concrete to the abstract, when discussing the film *The Wizard of Oz*:

- Referential meaning, referring to the facts of the film;
- Explicit meaning, referring to the point of the film;
- Implicit meaning, which becomes more abstract and goes beyond what is explicitly stated; and
- Symptomatic meaning, reflecting or commenting on held cultural values and beliefs.

"In short, films [or any other art objects] "have" meaning because we attribute meanings to them.

> … Our minds will probe an artwork for significance at several levels, seeking referential meanings, explicit meanings, implicit meanings, and symptomatic meanings." (Bordwell 2001).

Any object, be that a word, an image, a held belief, a sound, and so on, can traverse any of these meanings and new meanings are accumulated as objects gain new contexts. To take Barthes' telephone example further, mobile phones have had many meanings associated with them, ranging from status icon to basic/essential commodity. The metamorphosis of an object is well shown in the film *Phone Booth* (Cohen 2002), where telephones gradually become a pivotal focus for social ethics.

It may seem that telephones hold a special place in this discussion; they provide an example of an object with many meanings and values. It is possible to see a telephone as indicating: intimacy, absence, distance, place, status, or a particular social status, and so on. It has the ability to simultaneously fill all of these functions depending on the focus of the viewer. When using such an object in an artwork the art maker guides that focus for the viewer.

When developing the semantic zones within which to improvise, pools of images and sounds, each containing similar or related meanings, are created. For example: Pool One may contain images of living mobile things and the sounds that are associated with them; Pool Two may contain immobile living things and their associated sounds; Pool Three may contain mobile non-living things; and Pool Four may contain immobile non-living things. From these pools the improvisers can select certain sounds and images from which to develop a work.

There may be subsets within each of these pools, in Pool One for example humans may be a subset; which may have further subsets such as: children or different ethnic/social groups. Pool Three can be seen similarly, for example cars may be a subset. Here there are a wide variety of "cars" available to choose from, each offering and eliciting a different resonance in the viewer. We can consider the different effects of a visual image of a 1949 Kombi compared to a current 7 series BMW; while both fit the semantic zones "Car", and its subset "German Car", each will cause different resonances in each viewer and consequently draw a very different response from the viewer. It is possible that the artist could predict, or assume, the resonances these objects may render in the viewer.

The improvisers select the images and sounds in performance through their improvisations. They also influence the presentation of the images and sounds, effecting elements such as cross fade times and methods, speed of playback, repetitions and so on. It is also possible for the improvisers to influence the results of each others input, where, for example, if one improviser triggers a certain event then the event selection that is available to the other improviser is adjusted.

The unpredictability of the results creates new paradigms through which the improvisers are forced to view their output while in the process of creating it. This requires addressing each co-incidence of image and sound anew, as each image or sound gains new contexts and therefore becomes understood as having different meanings.

Precedents

There are many precedents to this approach for this kind of collaborative improvised art making. Some practitioners, such as the HUB or Simulus, or the users within the DASE community, use computer technology extensively in making interactive improvised works. Computer based technology is fundamental to each of these collaborations, just as it is to the one discussed here.

One precedent that used no more elaborate technology than a pencil and paper, is the "*Exquisite Corpse*" game "played by several people, each of whom would write a phrase on a sheet of paper, fold the paper to conceal part of it, and pass it on to the next player for his contribution [resulting in sentences such as] "Le cadavre exquis boira le vin nouveau" (The exquisite corpse will drink the young wine)" [from where it reputedly got its name.

Figure 1. *Man Ray, Yves Tanguy, Joan Miro and Max Morise, 1929.*

The image above used a similar technique, where "Tanguy [drew the head], which dissolves in to a jungle scene by Max Morise, returning to a female anatomy schematically indicated by Miró, and terminating in "legs" in the form of a fishtail and an engineer's triangle by Man Ray." (exquisitecorpse.com)

This approach builds narratives, which may seem as nonsensical as Noam Chomsky's "Colorless (sic) green ideas sleep furiously", but still have the ability to draw considered responses. The statement "The exquisite corpse will drink the young wine" definitely conjures up an image. When a visual image or sentence, such as the ones above, is presented, there is an attempt by the viewer to make sense of it, and it is in making that effort that pleasure, or some other kind of stimulation, may be generated.

Improvising system and parameters

The author's initial improvising system premiered at the "E.scape" performance in March of this year and included a MIDI-guitar linked to both a Max/MSP patch and a multi-timbrel sound module, as well as a vacuum-cleaner on which MIDI triggers had been built to control audio and video signals.

There were several performative aspects of the vacuum cleaner. A single extendable telescopic radio antenna that was able to be rotated, acted as a lever for sending pitch bend information, in this case altering the focus or blur of the projected image. “Rabbit ear” telescopic TV antenna were attached to the base of the vacuum cleaner, and worked by the performer completing a circuit, or triggering a MIDI note message, when touching both antennas. Completing this circuit then triggered the stepping through of different frames of an animated character or object. Touching both TV antennas could also be made to play random phrases and notes of MIDI piano.

There are several multipurpose buttons on top of the vacuum cleaner, functioning as MIDI keyboard notes. Several wires coming out of the vacuum cleaner attach to other objects chosen for their semantic value as much as their performative potential. For example, pressing down on a stapler stopped any audio coming from this system. Sugar and salt canisters were wired up with buttons under their lids. Striking the lids of these containers with any object or hand added video sugar or salt to the onscreen mix, producing “eye candy” particle effects such as sprinklings or glowing stars or grains of falling salt. Two cymbal stands held a cymbal and circular saw blade, both of which had microphones directly underneath them.

These two devices were chosen for their direct physical contact as well as their percussive value. The audio signal received from their microphones was set up through the Max/MSP patch so that if they were played loud enough they would create visual changes, such as switching video channels or adding effects.

The vacuum cleaner did not replace the computer keyboard for all controls of video and sound for one performer. From the computer keyboard, several banks of video clips were arranged to represent different semantic zones. The first row of keys stored video clips relating to early parts of a narrative structure: scene setting, establishing beginnings of possible adventures and journeys. The second row of computer keys represented clips from the middle stage of a potential narrative film; development, challenges, obstacles and mysteries.

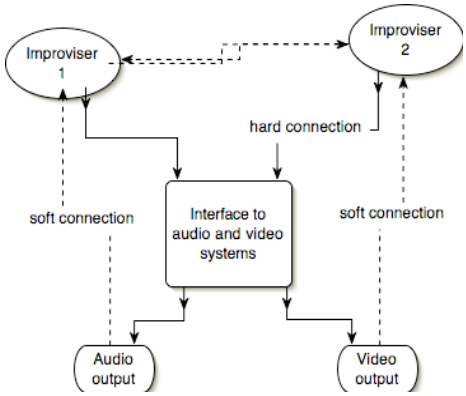
The third row of keys represented clips relating to resolution and a gradual return to the actual environment surrounding the performance. Each individual key on the keyboard could be repeatedly triggered to cycle through between 3 to 10 clip possibilities. In practice the performance involved an improvised weaving of clips from all sections.

The concept of this division of content between potential beginning, middle and end as semantic zones, was to provide the parallel of a possible chord sequence for improvising within. The intention was to provide the slightest trace of a narrative structure and in so doing allow the openness of abstraction and meditative reflection.

This process is being developed and refined to create a more directly interactive process, reflecting the goals given above regarding the influence the improvisers have over the materials available and how they are accessed.

Flow chart of information transfer in the improvisational process

The chart below shows the flow of information between the performers, the hardware, and the outputs. A continuous line denotes a hard connection, where the improvisers are directly connected via a cable, and a soft connection, denoted by a dotted line, where the improvisers are connected through interpreting each others input to the system and the resultant output, and their intuitive understanding of each other's personalities and approach to the improvising process.



This information flow is similar to that of any improvisation system, where there are direct and indirect connections between the improvisers, their instruments and their audience. The main differences here are that two modalities are being used, and that each improviser can fundamentally alter the instrument that they are playing and that of their collaborator(s).

This is a significant difference as it requires that all participants be responsive to the immediate situation, one in which basic, intuitive assumptions are continually changing or rendered irrelevant.

The result is a field in which to improvise that offers the participants reduced possibilities for assumptions or prediction. These form basic axioms around which improvisations, or any other temporally based art work, are measured by both the participants and their audience, at least at first exposure.

A possible example of this is where, for instance, pressing a key on a computer keyboard, or any other trigger, causes different results at each triggering, and at another time, cause the same result at each triggering. Obviously, this can lead to confusion for both audience and improviser if the paradigms that inform traditional improvisational settings are assumed.

However, if other improvisers collaborating in the work control these changes, this confusion can be mitigated as each collaborator develops an understanding of both the system, and the improvisational processes of each collaborator within that system.

Conclusion

By expanding the fields in which improvisation is occurring, in this case to include the visual and the aural, the paradigms that are associated with either of those forms are subverted, collapsed or system is gained, the im-

provisers' paradigms and shared processes will develop. It is important that these be subverted lest familiarity causes a regression to forms similar to those experienced when working in more traditional frameworks.

Here the more denotative and literal form of expressions of film are reassessed and reconstituted through a musical form, and musical forms are being equally reassessed and reconstituted through filmic form. Through this process, it may be possible that new paradigms can be developed.

References

Aksnes, H. *Meaning Generation in Music Listening* Unpublished paper, draft version.

Barthes, R. 1980. "Introduction to the Structural Analysis of Narrative" *A Roland Barthes Reader*. S. Sontag. London: Vintage. 251-295.

Berg, G. 2000. "Cognitive Development Through Narrative: Computer Interface Design for Educational Purposes" *Journal of Educational Multimedia and Hypermedia*. 9, 1. 3-17.

Bergqvist, S. Ekstrand, Martti. et al. 1993. *Revolver - Här är karusellen. Stig Bergqvist*. Martti Ekstrand, Jonas Odell and L. Ohlson.

Boden, M. 1995. "creativity and unpredictability (sic)" *Stanford Electronic Humanities Review*. 4, 2.

Bordwell, D. Thompson, Kristin. 2001. "Form and Meaning" *Film Art: An Introduction*. New York: McGraw Hill. 46-50.

Cohen, L. 2002. *Phone Booth*. J. Schumacher. www.exquisitecorpse.com (2005).

Miyazaki, H. 1988. *My Neighbor Totoro*. Tokyo: Studio Ghibli.