

Games, films and media literacy: frameworks for multimodal analysis

Andrew Burn
UCL Institute of Education

Please cite as:

Burn, A (2016) 'Games, films and media literacy: frameworks for multimodal analysis'. In Knobel, M and Lankshear, C (eds) *Researching New Literacies: Design, Theory, and Data in Sociocultural Investigation*. New York: Peter Lang.

In common sense terms, films, animation and computer games each have a singular, unified identity. This is true in popular discourse—but also it is often the case in academic film studies or game studies. Thus, it is common practice to maintain an idea of “the language of film,” as if it were a single mode. Similarly, we might talk over the last forty years or so of the developing “grammar” of computer games, especially pertinent for games scholars whose origins are in philology, linguistics or natural language programming in computer science.

Multimodal analysis places the emphasis rather differently on the wide range of modes these forms combine through digital media synthesized on single platforms both at the points of composition (e.g., digital editing software) and distribution (e.g., digital film projectors or videogame consoles). The object of the exercise is to discover what each of the modes (language, visual design, music, and so on) bring to the table, while also working out how the combined whole makes something we recognize as a film, or a videogame. Having done this, the researcher faces the really important “so what” question: what meanings are produced by these cultural productions, for whom, and in whose interest? What are the implications of such analysis for contemporary models of media literacy?

The approach adopted in this chapter is to consider the related forms of film, computer game and machinima as relatively settled cultural forms characterized by particular clusters of modes and media—dynamically multimodal, yet sufficiently stable for successive generations of designers and audiences to share some kinds of knowledge about “filming and editing” or “game engine.”

The approach to multimodality here is rooted in social semiotics. It emphasizes how moving image media use semiotic processes (organized sign systems) across different modes and media in order to fulfill particular social functions: to represent the world, to construct relations between social groups, to communicate particular interests, ideas, emotions.

Film: The kineikonic mode

The kineikonic mode is a designation for the moving image as a multimodal form (Burn, 2014; Burn & Parker, 2003). Kineikonic is a portmanteau of the Greek words for move and image: *kinein* and *eikon* respectively. While this theory draws on many of the classic conceptions of filming and editing as semiotic processes (e.g., Barthes, 1978; Bordwell & Thompson, 2001; Metz, 1974), it re-casts them as orchestrating modes: as framing structures which construct spatial and temporal shapes around the sounds and images produced by other modes (embodied action, speech, music, architecture, and so on), conceived of here as contributory modes.

It also presents the opportunity for a range of different analytical strategies. Three of these will be briefly outlined below with references to further reading; the third of these will be developed more fully as an example.

Analytical approach 1: Spatial and temporal syntagms

One approach possible through multimodal analysis is to disaggregate the spatial and temporal syntagms (a syntagm is a combination of signs, usually employing structures which make sense through a shared understanding between producer and audience, such as the grammar of language). In effect, the “eikonic” element of the kineikonic mode is the visual image.

Step 1: The “eikonic syntagm”

The first step proposed here is to isolate a single frame from the film being analyzed. Here, the structures identified by Gunther Kress and Theo van Leeuwen in *Reading Images* (1996) in relation to visual images apply. So the frame can be analyzed as a still image, identifying structures such as:

- Represented participants or characters, and the relationship between them in terms of implied action (realized as vectors), gaze, proximity and orientation to each other and the spectator, attributes (costume, hair, adornment, instruments)
- Location
- Lighting
- Construction of foreground and background, center and margin
- Construction of left and right, top and bottom
- Camera angle (vertical and horizontal)

Example 1. The example here is taken from Laurence Olivier's 1948 film, *Hamlet*. The scene shows Ophelia in her chamber. The scene represents her confusion about Hamlet's behavior, and how he apparently was succumbing to madness in the aftermath of his father's murder. Figure 9.1 shows a single frame, with Ophelia in the foreground and Hamlet emerging from an archway behind her.

The frame presents a two-part action sequence, the relation between the two characters set up by the forward movement of Hamlet (denoted by the feet and arms vectors) and the backward glance of Ophelia towards him. Ophelia's location in the foreground establishes a proximity with the viewer, echoed in the voiceover we will hear in the kinetic syntagm, in which she narrates the scene. We could go on to analyze the lighting, scenery, costumes, hairstyles and colors, and other features of the image. In general terms, the image presents uncertainty: the lighting is ambiguous, coming from the rear left and the foreground right; Ophelia's facial expression depicts apprehension in its slight frown (again complemented by the interrogative tonal contours of her voiceover).



Figure 9.1. A still from Olivier's *Hamlet* (1948).

Step 2: The “kinetic syntagm.” The second step proposed here is to analyze a segment of the temporal sequence containing the single frame analyzed in Step 1. Here, a critical division is made between the image track and the audio track. These contain a variety of distinct modes, which will be explored in more detail in Step 3. The analysis of the temporal sequence at this stage, however, might include:

- Action (the development of the action implied by the vectors in the still image)
- Music
- Speech
- Temporal sequence: the contraction of time (ellipsis), the expansion of time (stretch), the reordering of time (analepsis and prolepsis—flashback and flashforward)
- Camera movement (tracking, tilting, zooming)
- Editing (shot construction and transitions)

Example 2

Figure 9.2 shows a digital editing timeline containing this scene.

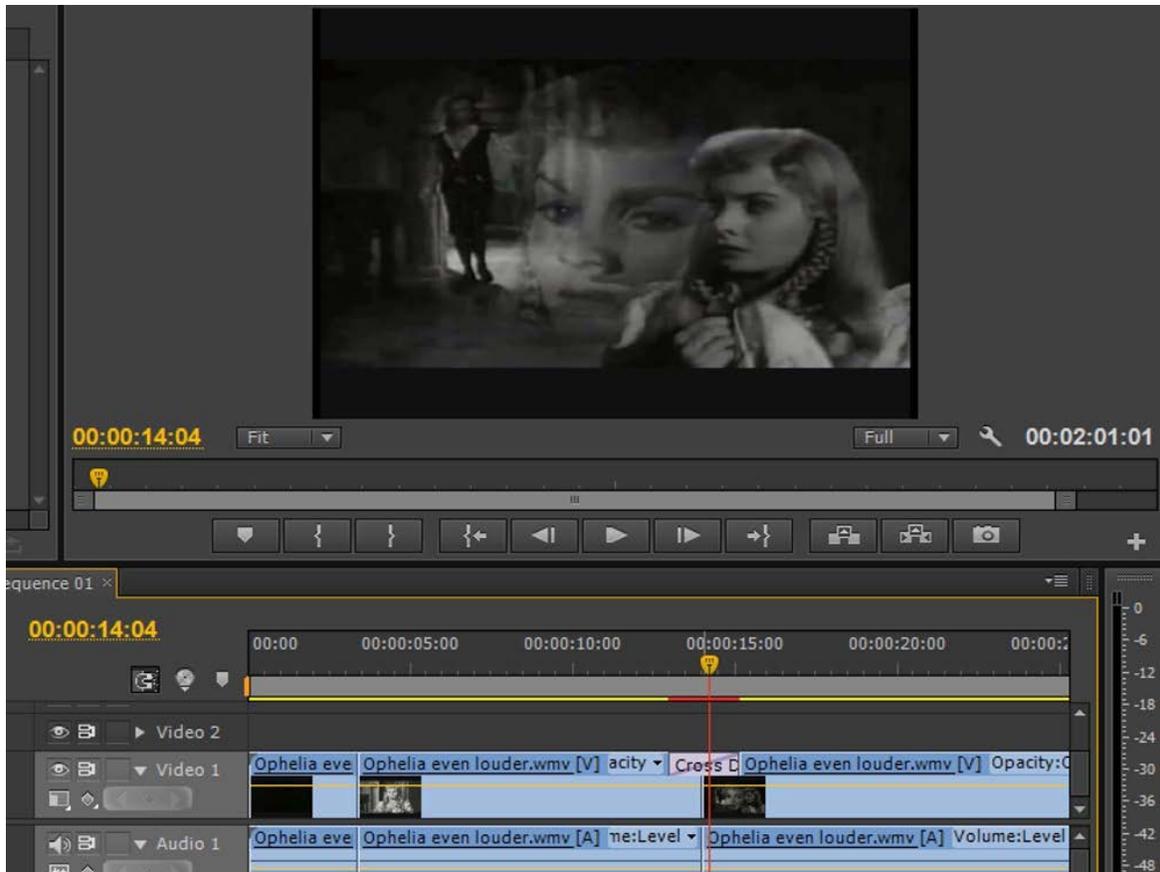


Figure 9.2. Digital editing timeline

The time indicator is paused on a dissolve effect, indicating a transition in the visual sequence. The transition is from a close-up of Ophelia's face in the center of the shot to the image represented in Figure 9.1, denoting a gradual shift from Ophelia as the only character in the scene to the two-character structure observed in Figure 9.1. However, as we add other features from the different modes into the analysis, the meanings become more complex, even contradictory. The effect of the dissolve can be considered, for example. What is its function? Why can the camera (or zoom) simply not move back gradually? In either case, camera move or shot change, the film is performing something not possible for the theatre-goer. Figure 9.1 might very well be an analysis of a moment from a stage-play: the camera-frame is performing a

similar function to the proscenium arch; the modes in play are all those that can be found in theatre. In the moving sequence, however, the mobile camera performs exactly those shifts in point-of-view that the theatre spectator in her fixed seat cannot. The move—especially the dissolve, which has no equivalent in theatre—draws attention to the cinematic nature of the sequence. We are reminded that this is a translation of Shakespeare into the popular medium of film, a move that was controversial at the time of its production. Indeed, in the accompanying book (Cross, 1948), Olivier and his editor are adamant that their intention is to preserve the dramatic and poetic qualities of the play as they see them. Yet this scene is a good example of the new meanings conferred by the new medium. Indeed, the entire scene is transformed: it never takes place in the play at all, except as a narrated memory told by Ophelia to her father. Film allows the voiceover convention, the temporal framing of the scene so that we understand it is a flashback, and further laminations of image and sound, which, though not impossible in theatre, would be less likely.

Analytical approach 2: The metamodal kineikonic

To look more closely at the interweaving of modes occurring here, this section presents a different analytical frame. *Metamodal* emphasizes the nesting of modes within modes in moving image texts. The prefix “meta-“ is used to indicate “beyondness” and “adjacency”—cultural forms and modes within, beyond and next to each other.

This model identifies *filming and editing* as orchestrating modes: the medium-specific meta-modes of the kineikonic. The process of filming and the process of editing effectively fall into two modes of design. *Filming* may resemble more the fluidity of speech; *editing* the fixity of writing. Filming produces spatial framing, angle, proximity and camera movement and

provisional duration, along with the orchestration of the “dramatic” modes: action, speech, set and costume. Editing produces temporal framing, and the orchestration of other contributory modes, especially sound, music and graphics.

In whichever way we approach the analysis of filming and editing, we must then consider the work of the contributory modes. This is not a new idea. For example, Roland Barthes, in his seminal work *Mythologies* (1957/1972), looks at a very specific aspect of Joseph Mankiewicz’s 1953 film of Julius Caesar: the meaning of beads of sweat on the brow of Marlon Brando. This reminds us that the great signifying systems of language, painting and music are not the only contributors to meaning: that a make-up artist dashing in to squeeze a drop of glycerin on Brando’s forehead is part of the complex of contributory modes.

Each of the contributory modes breaks down into more specific signifying systems. Language breaks into speech and writing, and each into lexis and grammar; each of these categories further decomposes into specific resources, so that speech can be analyzed syntactically, lexically, metrically, phonologically, and in terms of features such as tone and tonicity (Halliday, 1985).

Dramatic action breaks down into gesture, itself further sub-divisible into denotative, expressive, instrumental, symbolic and other kinds of movement (cf. Kendon, 2004) such as facial expression, proxemics.

Music can be analyzed in terms of melody, tonality, harmony, rhythm, style, instrumentation, dynamics and so on. As film theorists know, its function can be diegetic (music within the narrative) or non-diegetic (the more usual sonic coloring which we are enculturated into accepting as part of an otherwise naturalistic mode of narrative). While it is often explained as an

emotional accompaniment to film, it also has narrative properties, and, from a social semiotic point of view, carries particular social meanings (van Leeuwen, 1999).

A hierarchy can be constructed, then, from larger clusters of resources to minimal elements. The model shown in Figure 9.3 gives some idea of this, though it is indicative—many more semiotic resources are possible in individual moving image texts than can be represented here.

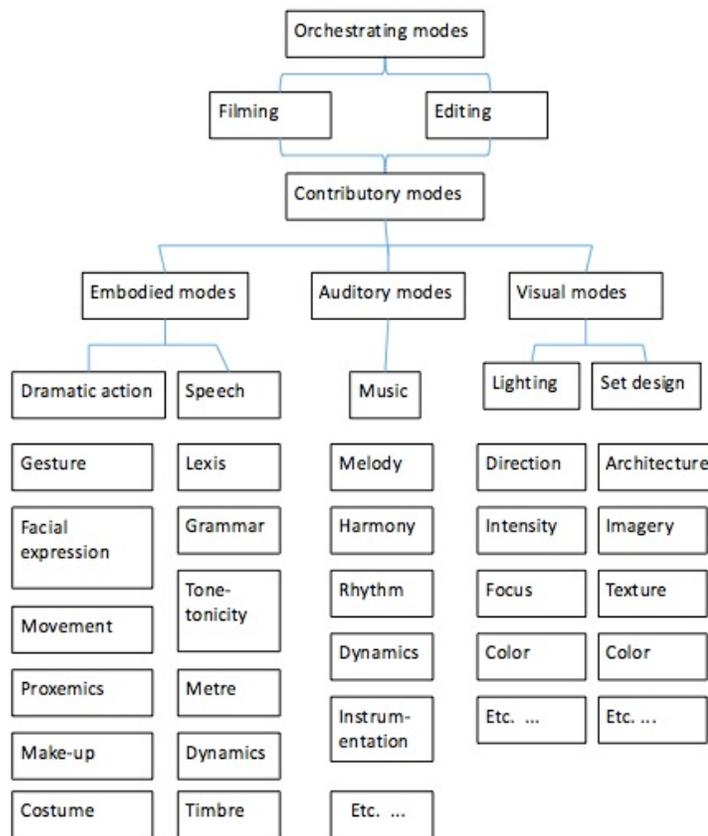


Figure 9.3. Metamodal structures

Step 1: Modal decomposition. One approach to the contributory modes, then, is decomposition: to break them down into progressively smaller resources, asking what the specialist role is of a

particular resource. This may be because the researcher has a particular interest in the contribution of one or two modes: the music, or the language, for example.

Step 2: Intermodal functions. However, the distinctive insight offered by a multimodal approach is to see how the modes work together. Decomposition may be a valuable analytical route for some researchers; and all analysts need to decide at what level of granularity they want to work; but many researchers will want to look across and between modes, asking how they connect to make meanings, and what semiotic principles work across them. Some of the features of individual modes identified in Figure 9.3 are in fact applicable across many modes, as van Leeuwen's work on rhythm in music, speech and editing amply demonstrates (van Leeuwen, 1985).

Example 3

Figure 9.4 shows a simple chart that transcribes the prominent modes in play in the *Hamlet* scene discussed earlier. The emphasis in this quite simple chart is on noting the modes and reading across them. No attempt is made to systematically represent space-time relations. The music column incorporates sections of William Walton's manuscript score for this scene.

Olivier was criticized for subordinating the high art of theatre to the popular form of cinema; but also for the opposite: producing cinema that was too "stagy." The effort to combine the forms can be seen in two decisions in this scene.

On the one hand, the scene is an example of Olivier's determination to fragment sequences of verse and drama as little as possible by cutting: it contains one of many very long shots in the film, which privilege dramatic continuity (in the same space and time) over

cinematic continuity (constructed by editing shots to create the illusion of continuous space and time). Here, then, filming as an orchestrating mode is used to promote the stage as the analogous orchestrating structure in theatre.

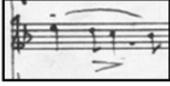
SPEECH		MUSIC	ACTION	FILMING	EDITING
As I was sowing in my Closet		Musical passage begins: D minor; low notes on 1 st violin.	Sewing; raises head	Medium C/U; central framing	First shot of the sequence fades from black.
		1 st violin passage continues.	Eyes move to L	Zoom to C/U	First shot
		1 st violin passage continues.		Lighting dim to diffuse low-key	First shot
Lord Hamlet with his doublet all vnlac'd,		Insistent repeated phrases.	Head turn to L; Hamlet stands	Superimposed shot backlit; main shot still low-key-lit	Superimposition of long-shot of Hamlet; Fast dissolve to next shot
Pale as his shirt,		 Falling phrase diminuendo	Hamlet walks forward	Wider two-shot (C/U and LS); key light raised; back and fill lights raised	Second shot continues ...

Figure 9.4. Transcription of scene from *Hamlet* (1948)

On the other hand, editing as orchestrating mode is used to provide a bracketing effect. Hamlet is introduced as a ghostly, superimposed figure (segment 4), suggesting both an otherworldly quality and that he is a product of her imagination or memory.

The introduction of the superimposed shot disrupts the apparent naturalism and theatrical continuity. Cook argues that this film, like the other three adaptations of *Hamlet* he considers, is firmly located in the mainstream “continuity” style of Western cinema, whose controlling aim is

to efface the markers of fragmented space and time to produce the illusion of spatio-temporal continuity (Cook, 2012). The superimposition in this scene, however, disrupts this, briefly disorienting the viewer: is Hamlet in her mind, or in her chamber? And, in relation to constructions of cultural value in film aesthetics, is this a formulaic popular piece or an example of radical montage? The uncertainty testifies to the confusion felt at the time about how to view this piece of popular Shakespeare; but also indicates the fragility of such reductive cultural valuations.

Even this brief transcription draws attention to meanings beyond the scope of this chapter. For a fuller analysis of other modes and their interaction, including the language and verse of the dialogue, the music, and the embodied performance of the two actors, see Burn, 2014.

Multimodality in Video Games

The cultural forms of theatre, cinema and games are closely related. They all, for example, involve the depiction of imaginary worlds, narratives made up of dramatic action and roleplay, the combination of spatial and temporal structures of signification. They can be seen in terms of adaptation theory, which considers how narratives which might begin, for example, in literature, migrate across other media, and how they are transformed in the process.

In this case, the example used will be a sequence from a videogame adapted from Shakespeare's *Macbeth* by two 13-year-old girls. Before looking in detail at this sequence, it is worth briefly noting some of the possibilities for multimodal analysis of games.

Analytical approach 1: Play, avatars and person

In adventure games and role-playing games, the player typically adopts the character of the protagonist—the player-character, sometimes referred to as the avatar. This is a good example of how the system of “person” which operates in the mode of language works differently in a computer game. The game presents itself in a kind of sustained second person, working through explicit or implicit offers and demands (you are in a gloomy dungeon; find the lost prisoner; will you take the left-hand corridor or the right-hand one?). The player is caught between a sense of first person narrative (*I am Harry Potter/James Bond/Gandalf*), and can operate some functions in this mode; and a sense of third person, observing the avatar from outside the game, and not in control of some of its functions. For extended discussion of this in relation to Japanese role-playing games, see Carr et al. (2006).

Analytical approach 2: Modality in Games

Like other moving image media, games often present a more or less naturalistic representation of a narrativized world, which may have cities, trees, lakes, people, and so on. Part of what makes it appealing is this naturalistic modality (Kress & van Leeuwen, 1996). Modality here (not to be confused with *multimodality*), means the claims made by the text to be true, authentic, valid, or credible, and how such claims are judged by the audience. The player may or may not find the streets, cars, water, and so on convincing—and as they explore the game, they will find other aspects of this naturalistic modality not available in films. They can, for example, experience of the physics of this world—if they throw an object at a wall, it bounces off. However, modality in games is complex. The naturalistic representation of the city might suddenly be interrupted by a screen full of statistics—of health points, inventory possessions, levels, scores and so on. This seems more like the kind of technological modality

identified by Kress and van Leeuwen with maps, charts and diagrams. But the player is of course not in the least disconcerted by this abrupt shift in modality. The interaction between this technological view of the game and the naturalistic view is learned within this genre, expected, and part of what makes the game authentic, believable, true to its genre. For an extended analysis of these kinds of modality in games, see Andrew Burn and David Parker (2003, Chapter 4).

Analytical approach 3: Coherence and cohesion in videogames

These features of textual organization—how the text makes sense through the structures that ‘glue’ it together—share some characteristics of coherence in language and other modes. The categories of cohesion identified by Michael Halliday and Ruqaiya Hasan (1976) are present, but require some adaptation. Their category of “reference,” for example, operates across modes, as in any multimodal text. So a magic amulet might be represented by a phrase in a text box, by a character’s speech, by an image, or by the 3D visual object itself; and the cohesion between these modes produces the aggregate meaning of the magic amulet. However, in the game, a further tie is present if the object is programmed to function in a particular way. If the player clicks on the amulet and it transports her to another place, then this additional cohesive tie, created through the procedural authoring of the game, provides an additional dimension, meaning and function. For a detailed discussion of these forms of cohesion in videogame design, see Burn (2008).

Analytical approach 4: Code as mode

This approach looks at game design, rather than at the game as a played text. While many aspects of games can be analyzed in ways similar to those applied to film, looking at functions of

the moving image, at systems of representation, and at structures such as point-of-view, the means by which the game is created requires another analytical move. Unlike literature, films and theatre, videogames are computer programs, which work because they have been written as code. We need, as it were, to look under the bonnet of the car to see how this “engine” is put together (and the program running a game is referred to as a game engine). As Michael Mateas argues:

Code is a kind of writing; just as literary scholars would not dream of reading translated glosses of work instead of reading the full work in its original language, so new media scholars must read code, not just at the simple level of primitive operations and control flow, but at the level of the procedural rhetoric, aesthetics and poetics encoded in a work (Mateas, 2007, n.p.)

The programming of games occurs at different levels, however. At the root of any computer program is binary code, but this is not what programmers typically work in. They will use computer “languages” which make it easier and quicker to give the computer the necessary instructions. The more abstracted these languages become from the numerical base, the more “higher-level” they are said to be. In the example below, the system for creating the rules which instruct the computer how to behave is not strictly a computer language at all, but a simple scripting system called a “rule editor.” Children working with the game-authoring software construct three-part rules for every event in their game. A simple example might be “If the player | clicks the key | the door opens.” These rules work like computer programming—they give the computer instructions. They also operate as procedural authoring: they devise a narrative

sequence, which depends on the dynamic relationship between the human user and the program (Ferreira, 2015). The player may choose not to open the door, in which case the narrative will take a different turn. Or they may open the door, and the programmed sequence clicks into place.

The proposal here, then, is that “code”—whether it be binary code, a computer language, or a simple scripting tool like a rule editor—is a mode operating in videogames below the surface. It will have particular affordances, always related to the requirement that it resolve into numbers or on-off, true-false states. In the scripting tool used here, this is manifested in the conditionality of the rule—all rules are “if-clauses” in this particular case. The condition prescribed in the rule either will or will not be triggered. This translates into a very particular kind of narrative, characteristic of games: that is, the options available to the player.

Step 1: Identifying the code. The analyst applying this approach will need access to the design process. This may take different forms: observing the designers creating the game engine and its rules and algorithms; access to the design program itself; or the abstraction of the rules or a subset of them from the design process (as is the case in the example below). The analysis of advanced computer languages is beyond the scope of this chapter; rather, the example given will illustrate the broad principle of treating the code as a mode which orchestrates the other modes in play in the game (for a more detailed account of procedurality and multimodality, see Ferreira, 2015).

Step 2: identifying the orchestrating functions of the code. The game engine orchestrates everything in the game, so some broad principles are necessary for the analyst’s focus.

Ultimately this will derive from the research question: the analyst may be interested in how

much agency is afforded to the player; or in the behavior of the virtual world; or in the nature and meaning of combat sequences; or in the player's interactions with other players in avatar form, or with non-player characters (NPCs), which are artificial intelligence entities functioning as characters in the narrative.

One focus might be derived from theories of play and game. One of the best-known and most-frequently-used models is the one elaborated by Roger Caillois in his book *Man, Play and Games* (2001/1958). Caillois presents six categories of play: Agon (competition); Alea (chance); Mimicry (role-play); Ilinx (vertigo); Ludus (rule-governed play); Paidia (loose, unstructured play).

So the analyst might be interested, for example, in the balance between Ludus and Paidia. Does the game tightly control the sequences of play—setting missions, for example, with clearly-defined win-lose states? Or does it produce looser structures in which the player is free to wander, explore, socialize? So the analysis of the game engine would focus on how such states are prescribed. There might be a sequence of code, for example, which prescribes how much damage can be received by the player when attacked by an enemy within a mission, and what resources are made available for retaliation. Or there might be a sequence which prescribes how the game-world behaves in relation to how the player can move, what actions they can perform, what sensory experiences they can access, and so on.

Step 3: relating the orchestrating and contributory modes. The analysis of the game engine can be related then to examples of play sequences. All media texts (indeed all forms of communication) present potential meanings, which wait to be actualized by a respondent's engagement. It is easy to forget this important point in the detailed business of textual analysis—

but games remind us of it forcefully. A game engine is perhaps the most extreme version of a set of potential meanings, since the variations possible in play are limitless—not only in terms of interpretation, as is the case with literature or film; but in terms of the actual structure of the narrative sequence. Even where the narrative is tightly confined and prescriptive (often the case with games adapted from literature or film), the player may decide how long to take over certain events, where to look, how often to switch between play mode and statistics screen, and so on. So to explore the relation between the orchestrating mode of the code and the contributory modes of visual design, music, speech, virtual embodiment and so on, a play sequence must be constructed and captured (for example by screen capture software) for analysis.

Example 1

This example is taken from a game made by two 13-year-old girls, in a project with Shakespeare's Globe to develop a game-authoring tool for Shakespeare plays based on the *Missionmaker* software. *Missionmaker* was developed by Immersive Education, and is now owned by MAGiCAL Projects at the University College London Institute of Education. The collaboration with Shakespeare's Globe, "Playing Shakespeare," was funded by the U.K.'s Arts and Humanities Research Council in 2012.

In the sequence discussed here, the player, as Macbeth, is able to kill King Duncan in his bedchamber. The task of the girls is to transform this event in the play into a game: to design a sequence in which the player has choices, multiple routes to follow, challenges to face, resources to meet them, and a win-lose state to complete the sequence. To do this, they create environments, characters, props and so on, using pre-rendered assets available within the software. To some extent, we can use similar analytical methods to those outlined above for

moving image texts. Games, at least graphical narrative games, share many properties with film: camera angle, framing, set, lighting, character behaviors and speech and so on. There are differences, too, as noted in the brief suggestions made earlier in this section: the relation between the player and the characters is different, for example, so that the girls are really designing a kind of second person offer, and a first-person character. These are profound differences in the nature of games when compared to films.

Figure 9.5 below shows the design screen of the two girls' *Macbeth* game. The large window shows the game world in design state, with the King Duncan character in the background. The transparent cylinder is a "trigger volume" invisible to the player—it is a programmed space, with a rule attached that specifies: "If the Player | enters the trigger volume | King Duncan wakes up."

The Rule Editor appears on the bottom left of the image. The rule active in the window specifies: "If Crown's state | becomes "owned by Player" | Player gets 600 Economy A points". The girls have programmed Economy A to be "Ambition," so that the player gains Ambition points, represented on a colored meter in the player interface, upon accomplishing particular tasks (an example of the "technological modality" described above).

The key characteristic of this procedural design is that it must produce conditionality—all programmed events depend on player interaction. This produces a particular kind of procedural narrative typical of videogames, and it can be analyzed in relation to the meanings of the game. In this case, for example, the player might choose to seize the crown after killing Duncan, eager to fulfill the ambition of the character, and hoping thereby to set in train a sequence of events leading to kingship—which could indeed be further programmed consequences in this case. Alternatively, the player might decide that they are overcome with remorse, and flee the scene,

ignoring such rewards, and setting a different chain of events in motion. These articulations of rule-governed behaviors and multiple routes through a text are characteristic of procedural authoring. As a matter of interest, some argue that such procedurality is not unique to videogames, only that they produce the effect through the automated processes of the computers, thus allowing a quantitative leap in what is possible (Bogost, 2006; Ferreira, 2015). A closely-related medium would be improvised theatre, in which actors might decide, from the resources available to them, to take one course of action or another (Boal, 1993).



Figure 9.5. The design interface of Missionmaker

The mode of code can be seen, as proposed in the case of filming and editing in the kineikonic mode, as the orchestrating mode of videogames. It creates the conditions which make the game possible; it determines how the game world and the characters within it can behave; it

prescribes the options available to the player, extending or limiting their agency; it decides what sounds will play and under what conditions. It can be analyzed as a kind of language, especially since it is commonly expressed as a kind of language. However, as in the case of film, it needs to be articulated, in the analysis, with the other modes and clusters of modes nested within the game design. Just as in film texts, all the historic modes of theatre (action, speech, setting, music, light) are nested within the orchestrating structures of filming and editing, so, too, in games all the modes of drama and the moving image are nested within the procedural logic of the game engine.

One example of this is the way in which the two girls have inserted the witches from Macbeth into their game level. In their first design session, they showed a particular interest in the witches, seeing these as appealing fantasy figures who we might trace back to a rich vein of fantasy tales from fairytales and films that they had experienced. In Shakespeare's text, of course, the witches do not appear in the scene that the girls were assigned to adapt as a game level. Their solution is to bring the witches back into the murder scene, where they work as a reminder to the player of the witches' earlier prophecies and their consequences. These witches also provide the girls with satisfaction through their uncanny presence, which they represent by a combination of three character models from the software: an old woman's head, a female superhero's torso and an alien creature's legs. The software allows the user to recombine the heads, legs and torsos of the character models, as in the popular children's picture books where similar effects are produced by flipping the three sections of the pages. The girls also attach particular special effects to these models: to one, an animation of fire; to another an animation of rain. These elemental attributes, which are not part of the Shakespeare play, nevertheless lend an archaic, magical quality to the characters.

In addition, the girls' design for the witches incorporates their own voices, making the witches speak prophecies as the models lip-synch the girls' speech. This is the only place in the game where the girls use their own voices. Elsewhere, for Lady Macbeth and King Duncan, they use the pre-recorded voices of professional actors. This choice suggests a particular motivation: to express empathy with the characters they are creating. Media literacy in this case is a multimodal combination of the virtual embodiment of the animated game characters with the actual voices of the girls as a dramatic performance, and the rules that the girls create to trigger the witches' speeches when a player walks into the invisible trigger volume. The media-literate practices consist of an association between the aspects of real and virtual embodiment (image and voice), and a subjective engagement with the fiction created by the teenage authors of the game and its players.

Machinima

Machinima is a portmanteau word combining machine and cinema, implying animation and anime. It is defined by Kelland et al. as "the art of making animated films within a realtime 3-D environment (2005, p. 10). It can be thought of as animation made from the 3D environments and animated characters of computer games or virtual immersive worlds. The first machinima films were produced by players of the game, *Quake*, in the mid 1990s.

Analysis here must pay attention to the nature of the virtual world from which the machinima text was created: its construction, behaviors, affordances and cultures. While in many

ways machinima resembles any other film or animation, the elements within it are differently created.

The characters, for example, are digital characters made possible by the virtual world or software, and typically have a set of specific shapes that can then be customized by the player. In this sense, they are more like actors dressing up for a role, with costumes, wigs, ornamentation, make-up and so on, than like animated characters which are drawn by the animator. Similarly, their movement is nothing like the frame-based motion created in traditional stop-frame animation. Instead, they will employ the animations provided by the software, allowing them to walk or run, dance or fight, fly or swim.

The metaphor adopted for this form of animation, then, is *lamination*. Lamination characterizes a particular thread in the history of the moving image, beginning with the superimposition of images through double exposure in photography, continuing through the forms of fantastic, magical animation exemplified by film-makers like Georges Méliès, and now evident in the practices of cinema effects, game design and 3D animation.

Step 1: Identifying the laminates

The laminated layers are oriented to particular modes, so to identify them involves the kind of modal disaggregation employed in the first example in this chapter, most easily accomplished through a simple transcription table. This allows the analyst to select which modes and intermodal relations will best address the research questions.

The modes are layered over each other, as shown in Figure 9.6.

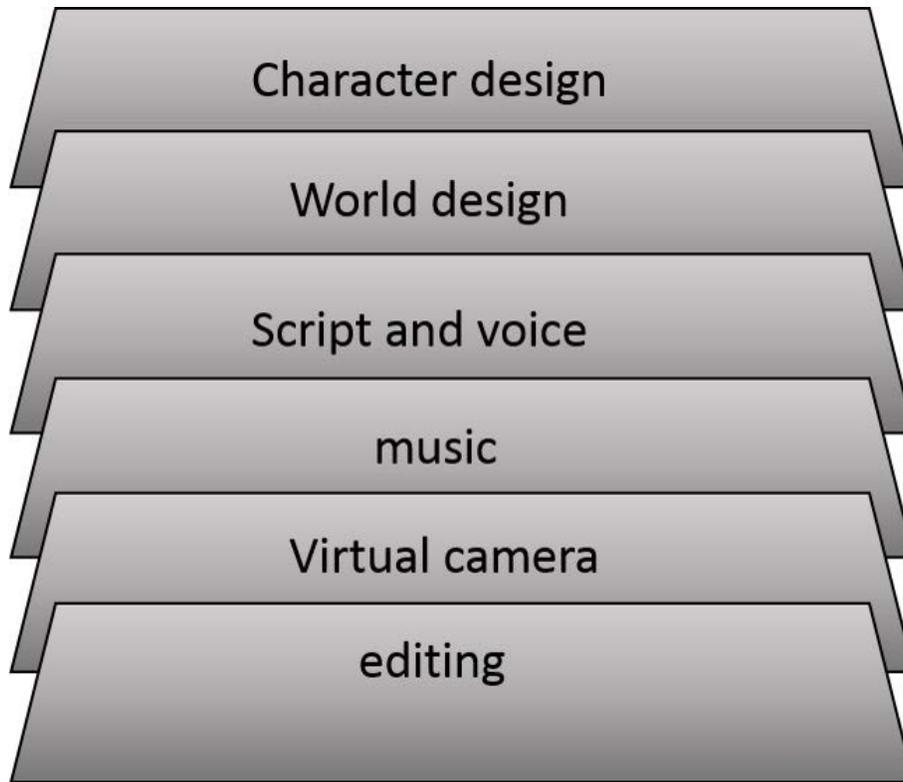


Figure 9.6. Modal laminates

So in any given sequence of the film, these laminates will appear. They are representational layers—layers of meaning—but also digital layers, which can be delaminated for editing (as appears in the example below); or for analysis.

Step 2: Drilling through the layers

Behind each layer, as it were, the production process exists in a further series of layers. To select Character design as an example, the machinimator will need to begin by selecting a character template, then customize it with clothing, make-up, tattoos, spectacles, and so on; then make it speak, employ whatever animations for facial expressions and action are available, film it with the virtual camera available in the software, and finally edit it in much the same way as any other

film. Figure 9.7 shows these production laminates. In selecting these and analyzing them in detail, the researcher may wish to explore the social motivations for the designs, and the meanings these contribute to the final animated film.

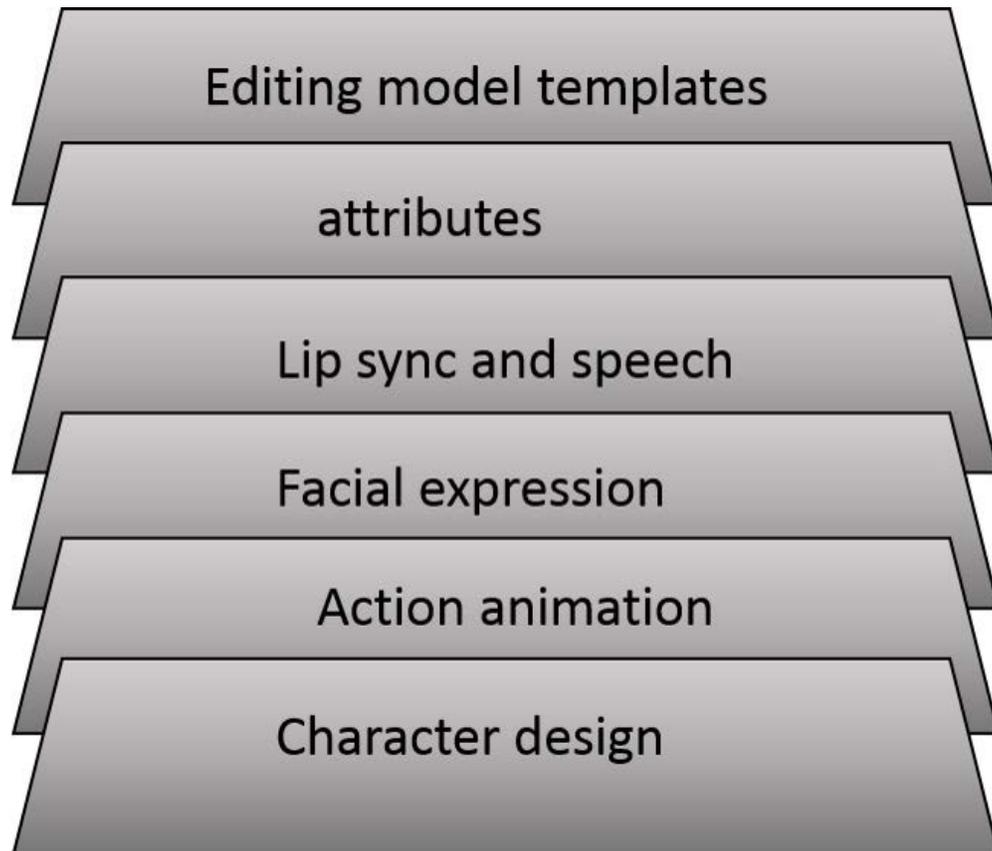


Figure 9.7. Production laminate

Example 1

This comes from an animated film made by a group of eleven-year-old children, *The Moonstone*. (This 2012 project, “Montage, Mash-up, Machinima,” was a collaboration with the British Film Institute and Moviestorm Ltd., and was funded by First Light). This animated film presents the tale of a teenage girl, Lily, whose parents were lost in a plane crash. She sets out to

find them with her grandmother and a strange Indian guide. He turns out to be a “good” werewolf, locked in combat with an evil werewolf. Lily helps him overcome his curse, beat the evil counterpart, and find her parents.

The children volunteered for four different design groups: set, character, voice acting, and music.

Finally, the class worked in pairs, each taking responsibility for a short section of the film. This involved drawing on the shared assets and using the 3D animation software, *Moviestorm*, to animate the characters through the action, attach dialogue, and construct shots and camera angles.

One instance from the making of *The Moonstone* film will serve as a detailed example of the nature of multimodal design in this project. It is a brief scene from the opening minute, in which the heroine, Lily, answers the door of her house to see a police officer who has the difficult task of breaking the news that her parents’ plane has crashed. The visual scene is designed by two girls, working with three sets of resources: the set design of Lily’s house and the street beyond it; the characters (Lily, her friend and the policeman); and the pre-recorded voice-acting of the children playing the parts of the three characters. Table 9.1 shows how the girls have constructed the sequence as four shots.

Table 9.1. Multimodal analysis of Lily and the Policeman

Shot	Framing and camera	Action	Speech	Gesture	Face	Music
------	--------------------	--------	--------	---------	------	-------

Shot	Framing and camera	Action	Speech	Gesture	Face	Music
1	Mid two-shot (Lily and friend) from behind Lily	Lily goes to open door		Door opens	(Back of head)	Guitar and bass loop
2	Close-up two-shot (Lily and officer) from behind Lily	Lily opens door, revealing officer	Officer: “Is one of you Lily Woods?” Lily: “Yes, that’s me”	Officer waves right arm	Officer’s face neutral	Solo violin descending minor arpeggios
3	Extreme Close-up (officer’s face); camera pulls back to close-up (head and shoulders)	Officer gives the news	Officer: “Well, um, I’ve, I’ve come to tell you that [pause] the plane that your mum and dad were on has crashed, or has gone missing – ”	Officer bows head and raises it again	Officer’s eyebrows raise in the centre, eyes widen.	Violin continues ...
4	Mid two-shot reaction (Lily and friend)	Lily weeps, comforted by friend	“– over the National forest”.	Lily covers face; she and friend turn to each other	Lily’s eyebrows raise; mouth turns down.	Violin continues ...

This transcription identifies the laminates and allows the first stage analysis of the modes in operation to be conducted. The next stage of analysis drills down into three of these laminates: speech, facial expression (character design), and music.

Laminate 1: Speech. The voice of the policeman is acted by a boy, who described in an interview the specific function of the role, and explained the emotional nature of his part:

It's a bit upsetting to tell someone that their parents might be dead...

And, drawing on his experience of an afterschool drama club, he describes the dramatic element in film:

Well in drama you have to use your face and your body, body language, and facial expressions: and in animated films you always also do that.

His job is that of the traditional voiceover artist. From Disney through to Pixar, voiceover acting is recorded in advance, and the animation made to fit the voice, rather than the other way round. In multimodal terms, the modes of voice and body action, unified in “live” drama, are segregated, and put back together in post-production. The voice retains its nature as dramatic performance, an embodied realization of character through intonation, dynamics of tempo and volume, and vocal timbre. By contrast, physical action, gesture and facial expression are designed visually as a simulation of performance. Nevertheless, the boy is aware of the relationship between the modes: that fictional narratives in film involve the same kind of embodied expression as drama. From a multimodal point of view, then, media literacy cannot be separated from the expressive modes of dramatic action and vocal performance.

These insights have a particular significance in the pedagogy of media education and the development of media literacy. The fragmentation of the curriculum into subject domains cuts

across the multimodal relationships of authentic cultural forms like film and games, and each domain tends to privilege its own modes. Media education (like film and media studies in Higher Education) may emphasize the orchestrating modes of filming and editing, and underemphasize important contributory modes such as dramatic action and voice. This constrained emphasis may occur in the day-to-day creative production work of teachers and children at least in part as a result of the conceptual frameworks they inherit from curriculum and examination designers. Nevertheless, it is not the case that specialist expertise should be abandoned. Film educators are necessarily expert in the art of filming and editing; music educators in melody and harmony; drama educators in embodied action and voice. It is a question of how each can become sufficiently aware of each other's domains, sufficiently collaborative at the right moment, and sufficiently creative in overcoming subject domain divisions.

Analysis of the boy's vocal performance shows that he delivers the lines with appropriate choice of intonation, volume and pace. Although the lines are scripted, he introduces three hesitation indicators: "um," a repetition of "I've", and a pause (Table 9.1, Shot 3)—an improvised performance of the anxiety implied by the narrative

Laminate 2: character design. Another student has been responsible for building Lily and the police officer, using the layers of composition in *Moviestorm* to construct face shape, hair, eyes, eyebrows, nose, mouth, and clothing. The job of the girls here is to build the moving image sequence on the editing track of the software: organizing it into shots, reframing the image with a virtual camera to change point-of-view, shot distance, angle, camera movement, deciding how to represent actions and speech. In many ways, this orchestrating work of the mode of editing is

similar to such work in conventional film. It involves, for example, learning the purpose and structure of shot-reverse-shot sequences to represent conversation (Burn, 2014).

However, the nature both of machinima and of this particular authoring tool, *Moviestorm*, produces some differences both from live action film and from conventional animation. The girls viewed the sequence in the rough edit of the whole film, and noticed that the policeman's expression did not convey the appropriate sentiment for his words. Aware of the mismatch between the modes, they amended the facial expression of the police officer, going into the relevant *Moviestorm* menus to find the slider tools to change the eyes, mouth and eyebrows:

when the policeman was telling Lily that her parents had gone missing then he was just like standing there with a really straight face and just saying sort of "I'm sorry your parents have gone missing" ... and then we had to, we went on the mood bit [makes gesture of sliding from left to right] and then we, cos there's only like a few bits like sad, angry, and happy, romantic and stuff, went on the advanced one and we made him more sad and more sort of, with the wiggly faced one.

In terms of the kineikonic mode, this piece of compositional redrafting is distinctively digital. It would not be possible in live action film without re-shooting part of the scene; nor in cel animation without redrawing the cels. Nor would it be possible in conventional machinima without re-enacting the sequence with the virtual actors in the virtual world. However, the machinima here is really a blend of machinima and 3D animation, and has the options of 3D animation to rework the character on the fly, digging through the laminates to the layers that need adjusting. In semiotic terms, it consists of reworking the kineikonic syntagm (the plane of

combination) through paradigmatic substitution (the plane of selection), making a closer intermodal match between the sorrowful intonation of the vocal track and the relevant parts of the facial expression. There is a selection of paradigmatic elements in play: replacing, in this child's words, "straight" with "sad", and working with a finer granularity of individual features. Such adjustments have similarities to the adjustment of lighting, audio volume, shot angle and so on in the semiotics of film, in that they are distributed across a numerical scale. The film semiotician Christian Metz argued that such elements could not properly be considered part of the "grammar" of film, since they were continuous scales rather than discrete paradigmatic elements (Metz, 1974). However, we may dispute this because of the perceptions of the human signmakers, and the ways in which they realize degrees of the scale in language. This produces something more like paradigmatic choices: camera angles, though adjustable on a continuous scale, become low shot, high shot, bird's-eye shot, and so on. Lighting becomes high or low, hard or soft, top or below, key, fill or back. Similarly, in these girls' choices, the slider scale filter for the policeman's eyebrows represents choices between neutral and sad.

Another point is worth mentioning. Since this is a digital medium, all choices at one level resolve into numerical ones. As Lev Manovich points out, this is one of the true distinguishing features of digital media (Manovich, 2001). However, once again, in terms of the interest of the signmaker (Kress & van Leeuwen, 1996), this numerical scale is transformed into paradigmatic choice. The child is interested in manipulating the representation of affect, and making a match between the modes of speech and virtual embodiment. In cultural terms, as noted above, we can see that the analogies are with similar activities in previously encountered virtual worlds which

children have engaged with, such as *The Sims*, and, as Jackie Marsh observes, a range of other online virtual worlds including *Club Penguin*, *Neopets*, *Moshi Monsters*, and *Habbo Hotel* (Marsh, 2014).

In the context of the history of the moving image, these practices evolve from a different source than the emergence of film from stage drama, or the conception of film as a documentary record. Rather, the culture and practice of animation in early cinema derives from traditions of illusion, magic, “trick” photography such as double exposure, and their associated practices of overlay, lamination, filtering and deceptive use of light.

In these children’s machinima film, these practices are still apparent. The appeal of animation is its readiness to represent the “impossible image:” fantasy elements of werewolf transformation, science fiction, time travel, and videogame play between virtual and “real” life.

Laminate 3: Music. The final layer to consider is the music layer. The music was composed by another group, using acoustic guitar, synthesized electric bass, and violin. The first sequence is a two-bar finger-picked ostinato, or repeated melodic pattern, with a bass line that suggests minor chords. The use of ostinato in film music can provide a dynamic, driving, insistent quality; or it can stabilize, even lead to monotony. Such ambiguity is well suited to this scene: the effect is of muted disquiet, appropriate for the dramatic mood of the piece, and anticipatory of the bad news to come (Figure 9.8).



Figure 9.8. The ostinato sequence accompanying Shot 1

The second sequence, which begins as the policeman starts speaking, is a soulful descending series of violin notes in a minor key. Minor keys traditionally signify sadness, while the melodic descent happens in steps: each bar descends, rises a little in the successive bar before descending further, and so on, so that an effect of a struggle against descent is conveyed. It has a baroque quality, accentuated by a trill on the penultimate note, a typical ornament of the baroque, suggestive of emotion (as in the music of Bach) (Figure 9.9). Here, the clear intention is to convey the tragic quality of the policeman's message.

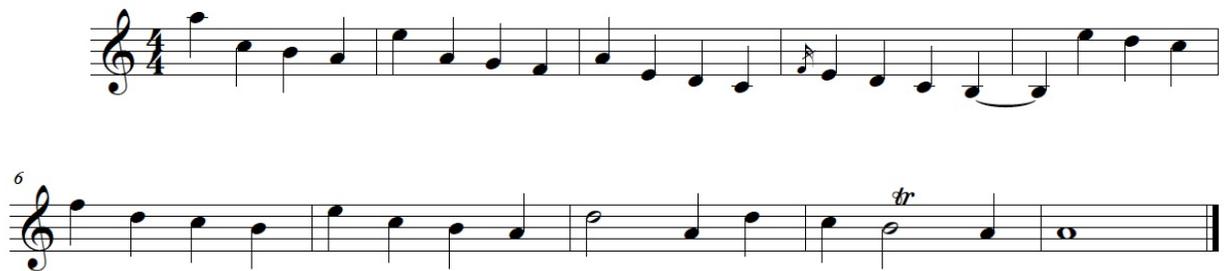


Figure 9.9. The solo violin sequence accompanying shots 2-4.

The music is performing a range of functions in this sequence. The loop structure of the ostinato in the first section is reminiscent of videogames and the unpredictable duration of their game-play sequences. It also indicates the tension between the girls' (i.e., Lily and her friend) ordinary evening at home and the growing anxiety caused by the parents' lateness.

The abrupt shift into the tragic, legato (smoothly-flowing) scales of the violin complements the policeman's message. The tonality of the violin, the downward trajectory of the arpeggios, and the minor key all work to emphasize the affective charge of the boy's performance, and the linguistic elements of the policeman's message.

This detailed analysis of the specific layers of vocal performance, visual design and music composition reveals two orientations in terms of the semiotics of the moving image and its "technologies of inscription," or material media (Kress & van Leeuwen, 1996). One orientation is towards the traditional lamination processes of animation, as discussed in the earlier sections of this article: the bifurcation of vocal and bodily performance; the layering of visual imagery to produce the "impossible image;" and the additional layer of musical signification, common in moving image culture since its inception. The other orientation is towards the "language of new media:" to the digital laminates of filters; the numerical scales of image adjustment and the new paradigmatic choices they produce; and to the playful compositional regimes of videogames and virtual worlds. We have also seen how this mix of practices requires the collaboration of educators with different specialisms across the arts: drama, literacy, music, media education, and ICT.

Conclusion: From Text to Context

This chapter has focused on the textual analysis of three quite different moving image texts. While each requires some degree of specialized analytical approach, they share many features, and thus invite a number of common principles of analysis. Most obviously, these are the ways in which they integrate the visual modes orientated to spatial logics, and the auditory modes orientated to temporal logics. This chapter has proposed in addition a metamodal approach, in which dominant processes of framing in space and time, realized in film as filming and editing, and in games as code, can be analyzed in relation to the modes they orchestrate.

The models proposed here also attend to the nature of digital composition: to the realization of modes as elements of digital editing timelines; to the digital code underlying game sequences; to the digital laminates overlaid on each other to create the actions and expressivity of machinima and 3D animation more generally.

These aspects of media texts also imply, as the game and machinima examples showed, an expanded model of media literacy and by the same token media education. Media literacy is multimodal, and cannot be constrained to a narrow focus on the orchestrating modes of filming and editing, but must connect with the modes of dramatic performance and computer code.

There remains a cautionary note. The texts explored here, like all text, remain systems of potential meaning until someone actually views or plays them. At this point, the research methodology must find ways to analyze the meanings made in this negotiation between film and spectator, game and player. Similarly, however, the production process cannot be fully deduced from the textual elements themselves. In the examples provided above, the researcher has talked to the producers or read their accounts to seek explanations and evidence of the sign-maker's motivation in their discourse. Methodologically, this requires some form of multimodal

discourse analysis beyond the scope of this chapter (e.g., Burn & Parker, 2003, Chapter 5).

Researchers in this field are urged to ask always what is significant about the meanings produced by the analysis; how they are situated in their socio-cultural context; whose interests they serve.

A simple but effective set of reminders, which offers an excellent ending point for any chapter on multimodal methodologies, can be found in the following design analysis framework proposed by Kalantzis and Cope:

1. Reference: What do the meanings refer to?
2. Interaction: How do the meanings connect the people who are communicating?
3. Composition: How does the overall meaning hold together?
4. Context: Where are the meanings situated?
5. Purpose: Whose interests do these meanings serve?

Kalantzis and Cope (2012: 200)

References

Barthes, R. (1957/1972). *Mythologies*, translated by Annette Lavers, New York: Hill and Wang

Barthes, R. (1978). *Image-music-text* (trans. S. Heath). New York: Hill and Wang.

Boal, A. (1993). *Theatre of the oppressed*. New York: Theatre Communications Group.

Bogost, I. (2006). *Unit operations: An approach to videogame criticism*. Cambridge MA: MIT Press.

Bordwell, D., & Thompson, K. (2001). *Film art: An introduction* (6th ed.). New York: McGraw-Hill.

- Burn, A. (2008). The Case of *Rebellion*: researching multimodal texts. In J. Coiro, M. Knobel, C. Lankshear, & D. Leu (Eds.), *The handbook of research in new literacies* (pp. 149-177). New York: Laurence Erlbaum.
- Burn, A. (2014). The kineikonic mode: Towards a multimodal approach to moving image media. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 375-385). (2nd ed.) London: Routledge.
- Burn, A., & Durran, J. (2013). Playing Shakespeare: Macbeth—narrative, drama, game. *Teaching English*, 1, 1-4.
- Burn, A., & Parker, D. (2003). *Analysing media texts*. London: Continuum
- Caillois, R. (2001/1958). *Man, play and games* (M. Barash, Trans.). Champaign: University of Illinois Press.
- Carr, D., Buckingham, D., Burn, A., & Schott, G. (2006). *Computer games: Text, narrative, play*. Cambridge, UK: Polity.
- Cook, P. (2011). *Cinematic Hamlet: The films of Olivier, Zeffirelli, Branagh and Almereyda*. Athens, OH: Ohio University Press.
- Cross, B. (Ed.) (1948). *The film Hamlet*. London: Saturn Press.
- Ferreira, D. (2015). The meaning potential of procedurality: Initial considerations on procedural semiotic resources. Working paper. DARE, UCL Institute of Education. Available from: <http://www.7luas.com.br/all/research/articles/the-meaning-potential-of-procedurality/comment-page-1/>
- Halliday, M. (1985). *An introduction to functional grammar*. London: Arnold.
- Halliday, M., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Kalantzis, M., & Cope, B. (2012) *Literacies*. New York: Cambridge University Press.

- Kelland, M., Morris, M., & Lloyd, D. (2005). *Machinima*. London: Course Technology PTR.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, UK: Cambridge University Press.
- Kress, K. & van Leeuwen, T. (1996) *Reading images: The grammar of visual design*. London: Routledge.
- Manovich, L. (2001). *The language of new media*. Cambridge MA: MIT Press.
- Marsh, J. (2014). The relationship between online and offline play: Friendship and exclusion. In A. Burn & C. Richards (Eds.), *Children's Games in the New Media Age* (pp. 109-132). Farnham, UK: Ashgate.
- Mateas, M. (2007). *Procedural literacy: Educating the new media practitioner*. Pittsburgh, PA: ETC Press, Carnegie Mellon University. Available from:
<http://press.etc.cmu.edu/node/205>
-
- Metz, C. (1974). *Film language*. Chicago IL: Chicago University Press
- van Leeuwen, T. (1985). Rhythmic structure of the film text. In T. van Dijk (Ed.), *Discourse and communication* (pp. 216-232). Berlin: de Gruyter.
- van Leeuwen, T. (1999). *Speech, music, sound*. London: Macmillan