

Kathrin Fahlenbrach (Martin-Luther-Universität Halle, Germany)

The Emotional Design of Music Videos. Approaches to Audiovisual Metaphors.

1. Introductory Remarks

Music videos aim to attract the attention of their public as fast and as effectively as possible, more than other audiovisual genre in film and television. With their close synthesis of music, sound and images they achieve this even with the most basic forms of the audiovisual composition. Therefore they use the whole archive of audiovisual techniques, of current and historical forms in film, television, and video art (cf. Fahlenbrach 2002).

Starting with the music, music videos create an intense emotional design that relates diverse primary emotional signals, without allowing the spectator to deepen his emotional reactions. Normally, they reduce their emotional design in order to ‘trigger’ primary emotional scripts and mood.

Especially for analyzing music videos we need a model that could describe how they design, even with the most basic audiovisual forms, emotion cues that can be perceived pre-conceptually: that is, before we interpret them semiotically and symbolically. In the following article, I’m going to propose the notion of *audiovisual metaphors* for such an approach (cf. also Fahlenbrach 2004).

2. First premises for modeling audiovisual metaphors: The Theory of Cognitive Metaphors

In their theory on cognitive metaphors, George Lakoff and Mark Johnson showed how much our perception and our thinking is marked by preconceptual, sensorial and physical experiences. According to Lakoff / Johnson, we permanently activate image schemata and gestalt structures that genuinely rely on bodily experience. As I refer to some main assumptions of this theory in my considerations on audiovisual metaphors, I will shortly outline some of its main aspects.

Lakoff presupposes that knowledge and thinking are fundamentally organized in a network of “*idealized cognitive models*” (ICM), mental models, that structure the background assumptions of our perceptions. The *idealized cognitive models* embrace several principles, forming together a structured whole, respectively a *gestalt*. He mentions four types of these structuring principles: *propositional structure*, *image-schematic structure*, *metaphoric mappings*, and *metonymic mappings* (Lakoff 1987, 68). According to this model, meaning is constructed

within a mental space that is structured as a network of schemata, building semantic links, relations, and nodes.

As I will demonstrate later, it is not only this idea of cognitive semantic networks that seems to be highly useful to describe audiovisual semantics, but also the emphasis on the visual and bodily structure of meaning. Following this model, visually and bodily based experiences are not only ‘mapped’ onto the semiotic representation of meaning, but knowledge and thinking are *themselves* deeply structured visually and bodily. Nevertheless Lakoff / Johnson draw a distinction between these two aspects, when differentiating between *image schemata* and *metaphorical projection*.

According to Johnson, image schemata are “those recurring structures of, or in, our perceptual interactions, bodily experiences, and cognitive operations. These schematic structures have a relatively small number of parts or components that stand in very definite relations to one another. So, whenever a single schema is instantiated in a number of different experiences or images, these same parts and relations recur.” (Johnson 1987, 79).

The meaning of *balance*, for example, emerges for Johnson from the “bodily experience in which we orient ourselves within our environment. These experiences involve preconceptual schemata that are the basis for the meaning of balance” (Johnson 1987, 74). Thus the meaning of *balance* embraces the very primary kinesthetic experience of our body in the world, and the emotional feeling of harmony or disharmony. At the same time it can be used to describe abstract relationships.

Another schema that relies on the interaction with our environment is the schema of *force*: every interaction requires *force* – either as we act upon others or as we are acted upon by them (Johnson 1987, 42). Concrete gestalt schemata of *force* are: compulsion, counterforce, diversion, blockage etc. (Johnson 1987, 45) – Similar to *balance*, both the meaning and the gestalt structures of *force* are based on bodily experiences: they are repeatable patterns of experience that include motion, directness of action, degree of intensity, and structure of causal interaction (cf. Johnson 1987, 44). As we will see later, audiovisual media directly appeal to those mentally rooted visual structures in order to create a most effectively emotional design.

Whereas *image schemata* represent a main cognitive structure, *metaphorical projection* is a central mode of semantic processes in the mental space:

“A second, related type of embodied imaginative structure (...) is metaphor, conceived as a pervasive mode of understanding by which we project patterns from one domain of experience in order to structure another domain of a different kind. So conceived, metaphor is not merely a linguistic mode of expression; rather, it is one of the chief cognitive structures by which we are able to have coherent, ordered experiences that we can reason about and make sense of.” (Johnson, xiv / xv)

On the one hand, bodily-based structures as *image schemata* are *metaphorically projected* from one experiential domain onto another. As I just mentioned, the bodily based meaning of *balance* e.g. can be metaphorically projected on emotional or abstract meanings of *balance*. On the other hand, Johnson emphasizes that we only perceive and understand our world in a unitary manner by integrating metaphorically different information und data, which are processed by the different senses. Namely the visual, acoustic, tactile and other stimuli can only be integrated to a cohesive perception, because they are metaphorically projected on to one another. As I will demonstrate in the following, this synthesis of different stimuli is fundamentally caused by *cross-modal perception*. Therefore on the most primary level cognitive metaphors constitute the very basis of our perception and understanding.

3. Audiovisual Metaphors in Music Videos

3.1 *The emotional structure of audiovisual media*

What implications could these conclusions have for audiovisual analysis? First of all, the model of cognitive metaphors demonstrates that audiovisual codes, as complex semiotic and symbolical forms, deeply rely on bodily based image schemata. This can't be overestimated for audiovisual analysis Because it means, that we always perceive first of all *physically* the aesthetic forms of films, television, and music videos (cf. also Tikka 2003). This is especially important when we look at the emotional design of audiovisual media, since those primary schemata essentially influence the emotional experience.

Before developing the notion of audiovisual metaphors, I will shortly outline in the following some main cues of the emotional structure in music videos, referring to emotion theory and to studies of cognitive film theory that are becoming more and more concerned with emotions in audiovisual media in recent years (cf. e.g. Plantinga / Smith 1999, Grodal 2002, Smith 2003) .

The associative network of emotions

After long debates on the primacy of emotion *or* cognition in our perception and thinking at the beginning of emotion research, most researchers on emotions today agree, that the emotional experience relies both on cognitive and affective processes, including bodily states (cf. Bartsch / Hübner 2004). Thus emotions can't be regarded as an isolated domain of our experience, but as a multidimensional system of psychosomatic and cognitive structures, mechanisms and processes that not only guide emotional experience but also bodily experience and cognitive understanding.

Fischer / Shaver / Carnochan (1990) for example demonstrate in their model of emotion prototypes, that emotional experience embraces several experiential domains that are linked into “functional wholes”:

„Emotions are complex functional wholes including appraisals or appreciations, patterned physiological processes, action tendencies, subjective feelings, expressions, and instrumental behaviors.” (Fischer / Shaver / Carnochan 1990, 85)

Along with the assumption that emotional experience, as well as perception and understanding in general, occurs within a network of multidimensional experiences, the associative structure of emotions is getting more and more obvious. This means, that the physical experience of sensorial intensity for example (such as volume in speech) can activate subjective feelings and action tendencies as well as cognitive schemata that guide semantic understanding. Thus looking at the network-structure of emotions, we inevitably recognize that perception and cognition in general are structured as networks. From such a holistic point of view, experience and understanding are deeply guided by associative processes.

As Grodal (2002) shows, this assumption is especially important when we look at semantic and emotional structures in audiovisual media. He mentions that “the meanings of a given phenomenon depended(s) on its relations to a ‘web-like’ structure of associations. Association is a fundamental mental phenomenon implying that phenomenon A is somehow linked with phenomenon B, so that an activation of A is related to an activation of B” (Grodal 2002, 64). Grodal points three crucial mechanisms of mental associations that characterize the associative construction of meaning:

- *establishing connections*,
- *chunking, i.e. grouping* (making gestalts, structures etc.),
- *and labelling* (Grodal 2002, 65)

Thus the mental construction of meaning highly depends on the framing and the context of audiovisual stimuli as well as on emotional and physiological experiences both of the addresser and of the addressee. Therefore audiovisual meaning seems to be highly flexible and fluid.

Within the audiovisual network of sounds and images, some elements activate cognitively established schemata that are mostly recognized consciously; at the same time a lot of audiovisual elements activate associative links that initiate connections that guide our experience and understanding of media at a preconscious level (cf. also the PKS-model of Peter Wuss, Wuss 1999).

“When viewing a film we will therefore perceive it at a certain perceptual level as directed by viewer and/or addresser, programming, for example by means of narrative schema, framing, zooming and other indexical procedures indicating formats of attention; and at the same time, the viewing will activate networks of associations below the threshold of consciousness, and activate superior, ‘propositional / abstract’ frames and themes” (Grodal 2002, 66)

Looking at the emotional structure and design of audiovisual media, this primary level of audiovisual creation and perception is especially important: that is, the elements that are perceived first of all physically and sensorially.

The audiovisual design of emotional scripts in emotion markers and emotion cues

The associative structure of emotions is especially obvious when we look at *emotional scripts*. As associative schemata, *emotional scripts* are not only related with cognitive components, but also with schemata of bodily experiences.

According to Fischer / Shaver / Carnochan (1990) *emotional scripts* are cognitively generalized structures that include the knowledge on prototypical *emotion cues*, *patterns of emotional reactions*, and *strategies of emotional control*. They assume that every basic emotion consists of a script of “behavioral and social events for the best or most typical case of the emotion, the essence of the category” (Fischer / Shaver / Carnochan 1990, 92).

“In line with other researchers’ usage, the term ‘script’ refers to both the generic representation of an event and the plan that is used to enact the event. (...) The script structure of the prototype for each emotion specifies typical antecedents and responses, including behavioral, expressive, experiential, and cognitive components. For negative basic emotions, there are also typical self-control or coping-strategies.” (Fischer / Shaver / Carnochan 1990, 92)

It is important to highlight that emotional scripts are not stable patterns that only *represent* generic aspects of emotional experiences. They *are* generic structures, namely associative networks which highly depend on the situational context (Bartsch / Hübner 2004, 64-84). This means that it is not the whole network of diverse physical, affective, and cognitive components that is activated in a given situation, but only those components that mostly relate to the given context.

“Emotional scripts are therefore not identical with the linguistically coded texts, that experimental subjects write down when asked to describe typical episodes of anger, fear, sadness, joy or love. (...) Emotional scripts are activated primarily by sensorimotor processing. Drawing on the theory of flexible concepts of Mangold-Allwin (1993), we suppose that script activation is shaped by synergy effects between the processing of context information and the activation of stored script information (translation by Bartsch).” (Bartsch & Hübner 2004, 82)

Thus Bartsch / Hübner emphasize in their work on emotional communication, that *emotional scripts* are flexible structures that don’t represent stable meanings of primary emotions, but generate dynamic and flexible emotional meanings that rely on the whole range of our experi-

ences (Bartsch / Hübner 2004, 80-82). The interpretation of emotional signals as *emotion cues* (for example in interpersonal interaction) thus depends on the current mental constellation of the activated components within the emotional script. As a result, the meaning of emotional signals changes with every new activation of an emotional script.

It seems very obvious that this associative and generic structure of emotional scripts closely relates to the associative structure of audiovisual media that I described before.

Another aspect of emotional scripts, that directly refers to the notion of audiovisual metaphors is its sensorial aspect. First of all the expressive, the behavioral, and the experiential components of emotional scripts, that are mentioned by Fischer / Shaver / Carnochan, deeply rely on sensorial and bodily based knowledge. *Emotion cues* for example can only be recognized by their sensory qualities: looking at the wide-open eyes of my vis-à-vis, I *see* and I *feel* immediately his fright or fear. Likewise, the knowledge on emotional reactions is inseparably interlinked with physical, kinesthetic or sensory-motor scripts – that activate for example archaic patterns of flight. Emotional scripts thus cover both cognitive and physical elements that are perceived and processed simultaneously.

Looking at the emotional design of audiovisual media, emotional scripts thus seem to have two main interesting aspects: First, the associative and generic structure of emotional scripts could be regarded as a crucial junction for the emotional structure of audiovisual media. As audiovisual media construct meaning within sequences of sound and pictures, both the production and the reception of emotional signals is associative, flexible, and changes permanently in the course of a film. Second, as I will explore in the following, the integration of sensorial and bodily based patterns of experience could be regarded as a crucial component of emotional design in audiovisual media.

This is especially obvious when we look at music videos. Since music videos barely have the time to develop emotional plots, they accentuate normally more than films the sensory, physical elements of emotional scripts, first of all of *emotion cues*. Following the mood-cue-approach of Greg Smith (1999, 2003), who also starts from an associative model of emotions, I assume that the construction of *emotion cues* is crucial for the emotional design in films.

“I argue that the primary emotions effect of film is to create mood. Generating brief, intense emotions often requires an orienting state that asks us to interpret our surroundings in an emotional fashion. If we are in such an emotionally orienting state, we are much more likely to experience such emotion (...).” (Smith 2003, 42)

Especially music videos create permanently and repeatedly emotion markers and emotion cues to actuate and establish certain, but undirected moods that are the basis for concrete, directed emotions. Whereas emotion markers are “highly visible textual cues” (Smith 2003, 45) that are mostly recognized consciously by the viewer, emotion cues are described by Smith as

the smallest aesthetic units of film, namely the formal elements that are perceived first of all preconsciously and sensorially.

“Films call on emotion prototypes (nodes of association often containing scripts) when asking us to interpret a characters’ actions, given their narrative situations and their facial expressions. But emotion cues also provide the possibility of nonprototypical access to the emotions, and therefore they tend to be used redundantly so that they may more predictably gain access to the flexible emotion system” (Smith 2003, 47)

By these directed *emotion cues*, such as nonverbal emotional expressions, the undirected moods can attain a concrete emotional reference: the emotional mood can thus be anchored and differentiated in relation to a certain emotional *object* or *subject* in the film.

As a result, the following aspects of emotional scripts could stand in the core of the analysis of audiovisual metaphors:

- The representation of elements of emotional scripts in certain motifs, in particular concerning *emotions cues*, and
- The composition of sensory qualities of emotional scripts, again focusing on *emotion cues*, within the audiovisual construction.

3.2 The emotional design of audiovisual metaphors in music videos

After having sketched some main theoretical cues to describe the emotional design of audiovisual media, I will now explain in more detail, how audiovisual metaphors could be defined and used to analyze music videos.

The most basic hypothesis that guides my definition of audiovisual metaphors, concerns their status within the audiovisual production. Even if one only looks at the commercial interests of the producers of music videos, it goes without saying, that they are highly interested in manipulating the perception of the viewers. As Grodal points it, “priming and direction of attention as ‘misdirection’ represent one of the main tools for producers of fiction” (Grodal 2002, 68).

Thus it can be assumed that directors of music videos *intentionally* create audiovisual metaphors to influence the perception of and the emotional effect on their public. Consequently the creation of metaphorical structures in music videos, unlike their perception, happens in a controlled manner. Even if directors, editors and the other members of the production team (art directors, choreographers and stylists) often link the various dimensions of single audiovisual effects in a highly intuitive way, the basic effects that they intend to initiate are established in the audiovisual design and can thus be scrutinized.

The next premise for defining audiovisual metaphors, concerns the basic elements of audiovisual creation: sound and image.

Michel Chion convincingly shows in his studies on film sound, how filmmakers can use the acoustic channel, not only to guide the visual meaning, but also to merge image and sound into an audiovisual unit. This is what Chion calls “synchresis”, namely the “forging of an immediate and necessary relationship between something one sees and something one hears” (Chion 1994, 4). Thus he describes a very elementary process of audiovisual aesthetics, starting from the assumption that image and sound refer to different senses, which are nevertheless integrated to a unitary meaning, what he calls the “added value”. The integration of acoustic and visual meaning occurs, when filmmakers link acoustic and visual qualities, which are already loaded with meaning in our perception of them.

Thereby he shows how closely aesthetic production and aesthetic perception are connected in audiovisual media. There are two main conditions that structure the acoustic and the visual, and in particular the audiovisual perception. First, acoustic stimuli are perceived initially *temporally*, whereas visual stimuli are initially perceived *spatially* (cf. Chion 1994, 11). Second, acoustic processing is much faster than visual processing, because the gestalt patterns of acoustic stimuli are immediately recognized, whereas the visual gestalt patterns are perceived in a holistic manner, so that it takes more time to discern them. These conditions of acoustic and visual perception have a big influence on audiovisual aesthetics: as sound is processed faster than the image and as it is perceived temporally, it has a big influence on the perception of temporal structures in the images shown. Chion mentions three dominant ways of this temporalization of the image by sound:

- “The first is temporal animation of the image: to varying degrees, sound renders the perception of time in image as exact, detailed, concrete – or vague, fluctuating, broad.
- Second, sound endows shots with temporal linearization. (...) synchronous sound does impose a sense of succession.
- Third, sound vectorizes or dramatizes shots, orienting them toward a future, a goal, and a creation of a feeling of imminence and expectation.” (Chion 1994, 13/14)

Thus, as a first step, *audiovisual metaphors* could be defined as a metaphorical mode of the audiovisual synthesis, within which acoustic and visual schemata are projected onto one another in a metaphorical way. In the “synchresis” of image and sound, primary visual and acoustic gestalt patterns are related to an audiovisual network that is experienced first of all physically.

As described above, according to Lakoff / Johnson, the gestalts that are created in audio-visual media are much closely related to the image schemata and gestalt patterns that are cognitively established in our brain.

Referring to the schema of *balance* for example, a slanted camera angle and image composition can be related in a music video with break-beats and asynchronous sounds. Thus image and sound create together an audiovisual metaphor for ‘disharmony’ and the ‘loss of balance’. As a result it can implicate not only an emotional effect but also a more complex narrative meaning.

Cross-modal metaphors in audiovisual perception and creation

The metaphorical projection not only refers to *gestalt patterns* but also to other primary structures of perception that fundamentally guide cognitive understanding and emotional reactions. Thus the tuning of the single acoustic and visual elements is especially effective, if it is oriented toward the basic mechanism of our multi-sensory perception: that is the *cross-modal processing*. It is the main precondition for our unitary perception, since it allows the evaluation of every sensory stimuli upon the so-called *amodal qualities*. Though every sense processes different stimuli: acoustic, visual, tactile etc. data are assessed by *amodal qualities* such as *rhythm / duration, intensity, movement, number* and *gestalt* – qualities that can be processed by every sense (cf. Stern 1993). As a result, these amodal qualities integrate the different sensorial information to a unitary perception in our brain.

It seems very obvious that this process of cross-modal projection can be regarded as a metaphorical process. Since metaphoric understanding is “the ability to perceive similarity among seemingly dissimilar objects” (Cytowic 1993, 207), the mapping of different sensorial information onto amodal qualities as *intensity* or *rhythm*, is the most basic metaphorical process in our perception. Therefore neurologists as Lawrence Marks and Richard Cytowic speak of “cross-modal metaphors”. They assume, similar to Lakoff / Johnson, that our perception is not only fundamentally structured metaphorically through cross-modal processing, but that the resulting cross-modal metaphors are the basis of every semantic meaning in language and other semiotic systems.

“It may be through this perceptual, cross-modal equivalence with regard to intensity that people are capable to uniformly interpreting intensity in cross-modal metaphors: Because dim lights resemble soft sounds, the word “dim” connotes “soft”, and because bright lights resemble loud sounds, the word “bright” connotes “loud”. By analogy, “soft” implies “dim”, and “loud” implies “bright”. (...) Metaphoric translations of loudness to brightness and of brightness to loudness at the verbal level follow the rules of cross-modal perception.” (Marks / Hammeal 1987, 5)

As they reveal the sensorial, and thus the irrational aspects of our understanding and thinking, cross-modal metaphors are much closely related to emotional processes. Cytowic thus assumes that metaphors are the very basis of emotional perception.

“I mean that metaphor is experiential and visceral, an irrational transfer of connotations from one thing to another. The emotional, irrational self is wise beyond knowledge, and we can see this wisdom in the way metaphor physically encapsulates our relations with the world. While metaphor is a means of seeing the similar in the dissimilar, it is empathically not rational analysis.” (Cytowic 1993, 206).

This assumption, concerning the sensorial and bodily based processes of emotional perception, can be confirmed, if we look at emotional communication. Following studies on nonverbal interaction, the nonverbal representation and interpretation of amodal qualities seem to be crucial for emotional interactions like emotional assimilation, empathy, emotional contagion etc. (cf. Bartsch / Hübner 2004, Forgas 1994).

During nonverbal interaction we guide our attention to visual signals in mimicry and gestures of our vis-à-vis and coordinate them with the acoustic rhythm and intensity of his speech. Daniel Stern shows in his studies on the psychological development of infants, that mother and child develop common *gestalt patterns* in nonverbal interaction, to communicate their feelings and emotions (Stern 1993, Bartsch / Hübner 2004). These *gestalt patterns* include *rhythm*, *intensity*, and other amodal qualities of speech (such as loudness and pitch), facial expression, gestures etc. It can be assumed, that those primary gestalt patterns of emotional communication aren't reduced to mother-child-relations, but are developed and differentiated individually, socially – and by media.

Conclusion:

As we see, both perception and creation of audiovisual media seem to rely deeply on metaphorical processes. Before I will demonstrate the use of audiovisual metaphors in media analysis, using the example of a music video, I want to sum up some core points.

As I mentioned above, audiovisual metaphors can be defined as the metaphorical mapping of cognitive, emotional, and sensorial elements of images and sound into an audiovisual network, that is initially perceived bodily and emotionally. Looking at the emotional design, there are two main elements of audiovisual metaphors, that directly refer to metaphorical perception:

First, the audiovisual creation of *emotion cues*. They can be created both by the construction of certain motifs, that initiate prototypical emotional scripts and by the formal, aesthetical design of prototypical gestalt patterns of emotional scripts, that are perceived preconsciously. Thus audiovisual media are able to build audiovisual metaphors that directly refer to primary

gestalt patterns, cognitive schemata, and emotions cues, as they are generated in emotional scripts.

Second, the audiovisual *synchresis* of image and sound that is based on the metaphorical mapping of the acoustic and visual elements to a coherent audiovisual meaning. In order to reconstruct the acoustic and visual elements, that are forged within the audiovisual design, it seems highly useful to use amodal qualities, such as: *rhythm /duration, intensity, and gestalt*.

4. Audiovisual Metaphors of ‘fear’ in “Come to Daddy”

Now, I will demonstrate these core points with reference to the music video “Come to Daddy”, from Aphex Twin, directed by Chris Cunningham (1997).

It is a well-known concept video, that isn’t in fact representative of the genre, but it illustrates very vividly the possibilities of music videos to create emotional design. Furthermore, the video is very interesting, because it uses typical elements of horror films as an additional emotional dimension. I focus on the introductory sequence (1,27 minutes) and the first five seconds of the song.

Before the music starts, the following sequence is shown during the first minute: An elderly woman with a dog is trying to get to the entrance of a high rise building located in a very deserted urban wasteland: a dirty place in front of a car park. In the middle of the place we see a garbage heap with a broken television monitor. The woman and the dog, that are crossing the place, are distracted and harassed by the television monitor that is suddenly switched on and by a gang of aggressive, cloned children.

The emotional design of the clip is apparently directed towards emotional scripts that are closely related: *fear, anger* and *disgust*. I will reduce my analysis here on the design of *fear*. Therefore I like to refer to prototypical metaphorical attributions of *fear* that Zoltán Kövecses (2000) observed in large empirical studies. Kövecses considered, that *fear* is attributed mostly with linguistic expressions that are grouped around the following metaphors:

- Fear is fluid in a container,
- Fear is a hidden enemy,
- Fear is a supernatural being,
- Fear is an opponent in a struggle,
- Fear is insanity, is illness,
- Fear is a natural force,
- Fear is a burden,
- Fear is a social superior,

- The subject of fear is a divided self. (Kövecses 1999, 23)

These attributions refer to prototypical elements of the emotional script of *fear*. Thus they are related with prototypical emotions cues and gestalt patterns, which are at the same time distinctive and flexible: That means that they include primary prototypical elements that are specified differently in diverse contexts. Therefore, these attributions indicate prototypical semantic and formal characteristics for the representation of *fear*, as it is visualized and / or designed acoustically in audiovisual media. Since their semantic and formal appearance is at the same time flexible, they can be projected on diverse representations of *fear* in language and in audiovisual media.

It is amazing, how much these linguistic attributions to *fear*, that Kövecses observed, correspond with the audiovisual metaphors of *fear* in the music video. For one thing, the representation of threat and *fear* in certain motifs can be described accurately by these metaphors. First of all the description of the *subject* and the *object* of *fear* fits the mentioned prototypical metaphors: An old woman is the subject of fear. She is suffering from its burden. The video establishes first a hidden enemy that is getting more and more concrete and aggressive: we only see the gang of the cloned children towards the end of the sequence. Nevertheless, they are indicated by single visual and acoustic elements: the distant echoes of laughter, the sound and pictures of running steps, and the diffuse picture of persons, hidden in the shadow of the car park, observing the old woman on the place.

Beneath, the enemy is shown as a supernatural, insane opponent – represented both by the cloned monster kids with the smirking faces of Aphex Twin and by a monster in a tube: When the woman passes with her dog the television monitor, it is suddenly stimulated by a supernatural force: the monitor shows a monstrous face that is screaming at her and her dog. Thus in a first step we can discern in the audiovisual motifs the representation of certain emotions cues, that highly concur with the prototypical, metaphorical attributions of fear, explored by Kövecses.

But how are these motifs as metaphorical attributions of *fear* created in the formal audiovisual composition itself? The primary *gestalt structures*, which are shaped visually and acoustically, correspond here directly to the mentioned schema of *balance* and *force* that Mark Johnson analyzed. Also in the video, they are a main interface between the sensory gestalt of *fear* and its meaning.

With the audiovisual attribution of *force* and *balance*, the emotional design attains a sensory *gestalt* that also embraces other *amodal qualities*: *intensity* and *rhythm*. With *fear* e.g. as

a rising force, the *intensity* of the emotional design rises. Connected with the amount of *intensity* is the acceleration of the *rhythm* – the visual and the acoustic rhythm.

In the following, I will describe the concrete tuning of *gestalt*, *rhythm* and *intensity* in reference to the schema of *balance* and *force*, using the example of the introductory sequence.

Before we know the object and the subject of fear in the video, the first shots already establish with a dense confusion of image and sound an atmosphere of threat and fear.

Referring to the schema of *balance*, the metaphorical composition of image and sound indicates a growing loss of control, and thus, of balance. That refers first of all to the old woman as the subject of fear. But this ‘loss of control’ is also directed toward the disorientation and emotional disconcertion of the viewer.



At the beginning, the black masque in the first shot opens slowly, and shows, in a slightly swaying camera movement, a low angle shot of the building, accompanied by the slow rhythm of the sound. All this is accompanied by a dim, vibrating sound. Here, it is initially the visual perspectives that are getting out of balance. While the sound is vibrating in regular sequences, the images are slanting and swaying. But at the same time, the search for balance is mediated visually: in a short steady shot the oblique view at the buildings is put upright; thus the image accords to the impulse of an imaginative spectator to adjust cognitively the imbalance.

Regarding the schema of *force*, images and sound indicate the growing threat and the related rise of fear. During the first shots, two visual spaces are established: the ‘real’ scenery of the buildings with their surroundings and the ‘virtual’ space of the inserts. The scenery of the buildings is revealed from the perspective of a hidden protagonist, shown in a moving, subjective camera – later we can attribute this view to the old woman. In the ‘virtual’ space of the inserts a supernatural, mechanical force or power is established.

Both spaces are specified acoustically: The view of the buildings is accompanied by a persistent, buzzing sound at a middle pitch that together with the swaying angles and the diffuse colors of the scenery creates a hostile atmosphere. The virtual space, on the other side, is characterized acoustically by a bright sound that slowly rises to a vibrating chime.

Here, we can consider a direct reference to Kövec's leading metaphor of fear: ‘fear as fluid in a container’. In the insert we see splashes of water, which are, by a metallic noise, acoustically characterized as a sharp and dangerous force. The visual fluidity and the acoustic sharpness build here an ambiguous, suspenseful and irritating metaphor of *fear*.



Nevertheless, the object of *fear* stays hidden and diffuse in the ‘real scenery’ until this point: hidden by the subjective camera, hidden by the short over-the-shoulder-shot out of the car park, and hidden by the insert of the running feet of the kids and the combined echo of voices.





The anxious gazes of the woman are the first emotions cues up to this point that give a concrete reference to the diffuse mood of *fear* and threat. At the moment when the dog barks at the TV and the screaming face appears on the screen, the object of fear is concretely comprehensible. At this point, the music begins: a loud and musty sound that accompanies together with single echoes of fast metallic beats the distorted speech-song of the screaming monster face in the tube, always repeating the same refrain: “I want your soul!”



With the start of the music, the suspense is immediately broken: the latent threat turns into expressive aggression, the anxious mood turns into shock. The effect of this change is particularly strong, as the rhythm and the intensity of image and sound contrast radically with those of the introductory sequence. The dim sound turns in the loud, aggressive music that is accompanied by the noisy antics of the kids. The slow motion ends abruptly and camera moves, cuts and movements accelerate rapidly. Thus the analyzed emotional design of the introductory sequence creates an extremely intense state of suspense that prepares efficiently for the shocking effect of the music.

As I have shown, a dense network of emotional stimuli has been shaped, and it is difficult for the viewer to escape its effects. Because basic physical and cognitive experiences of *fear* were projected onto the visual and acoustic elements of the video, Cunningham created audiovisual metaphors that are experienced first of all bodily and pre-consciously. Due to their immediacy and emotional force audiovisual metaphors are especially attractive for the makers, and analysts, of music videos.

Literature

- Barcelona, Antoni, Ed. (2003). *Metaphor and metonymy at the crossroads. A cognitive perspective*. Berlin: Mouton de Gruyter.
- Bartsch, Anne / Hübner, Susanne (2004). *Emotionale Kommunikation. Ein integratives Modell*. Unpublished dissertation. Martin-Luther-Universität Halle.
- Chion, Michel (1994). *Audio-Vision. Sound on screen*. New York: Columbia UP.
- Cytowic, Richard (1993). *The man who tasted shapes. A bizarre medical mystery offers revolutionary insights into emotions, reasoning, and consciousness*. New York: G.P. Putnam's.
- Cytowic, Richard (1993). *The experience of metaphor*. In: Cytowic (1993), 206-10.
- Fahlenbrach, Kathrin (2002). *Feeling sounds. Emotional aspects of music videos*. In: *Proceedings of the 8th Conference of the International Society of Literature and Media*, August 2002, Pécs. <http://www.arts.ualberta.ca/igel/IGEL2002/Fahlenbrach.pdf>
- Fahlenbrach, Kathrin (2004). *The aesthetics of media perception as Aisthesis? Considerations on audiovisual metaphors*. In: *CLCWeb: Comparative Literature and Culture* <<http://clcwebjournal.lib.ourdue.edu/>>, Special Issue: *German Media Studies* (forthcoming)
- Fischer, Kurt W. / Shaver, Phillip R. / Carnochan, Peter (1990). *How emotions develop and how they organize development*. In: *Cognition and Emotion*, H. 4, 81-127.

- Flückiger, Barbara. Sound Design (2002). Die virtuelle Klangwelt des Films. Zürich: Schüren.
- Forgas, Joseph P. (1994). Soziale Interaktion und Kommunikation. Eine Einführung in die Sozialpsychologie. Weinheim: Beltz.
- Grodal, Torben (2002). Moving pictures. A new theory of film genres, feelings, and cognition: Oxford UP. (2. Edition)
- Johnson, Mark (1987). The body in the mind. The bodily basis of meaning, imagination, and reason. Chicago: UP
- Kövecses, Zoltàn (2000). Metaphor and emotion. Language, culture, and body in human feeling. Cambridge: UP.
- Lakoff, George (1987). Woman, fire, and dangerous things. What categories reveal about the mind. Chicago: UP.
- Marks, Lawrence / Hammeal, Robin (1987). Perceiving similarity and comprehending metaphor. Chicago: UP.
- Plantinga, Carl / Smith, Greg M., Eds. (1999). Passionate views. Film, cognition and emotion. Baltimore: John Hopkins UP.
- Smith, Greg M. (1999). Local emotions, global moods, and film structure. In: Plantinga / Smith, ed. 103-127.
- Smith, Greg M. (2003). Film structure and the emotion system. Cambridge: UP
- Stern, Daniel (1993). The interpersonal world of the infant. London: Karnac Books.
- Tikka, Pia (2003). (Interactive) Cinema as a Model of Mind” in connection with the presentation ”Cooking in the Crucible” by Crucible Studio, University of Art and Design, Helsinki. In: *Digital Creativity Journal* (ed. Maureen Thomas, CUMIS), fall 2003.
- Wuss, Peter 1999. Filmanalyse und Psychologie: Strukturen des Films im Wahrnehmungsprozess. Berlin: Edition Sigma (2. Edition)

mail-address:

fahlenbrach@medienkomm.uni-halle.de