Structural relationships of music and images in music video

ALF BJÖRNBERG

Introduction

In the course of the last decade, the body of writing on music video has grown to sizeable proportions. The reason for the present addition to this bulk of literature, in spite of the subject seemingly approaching the state of exhaustion, is that musical semiotics are still rarely applied to the field. It is a fact that pop and rock music have always been heavily infused with socially determined meaning such that an autonomous musical aesthetics appears clearly insufficient to explain their significance; however, to what extent and how this significance is linked in with particular musical structures as such is still largely uninvestigated. In my view, music video may perhaps be less interesting as a phenomenon in itself than as source material for an 'empirical semiotics' of popular music, shedding light on signification processes of a more general applicability. Furthermore, the distinctive features of music video may arguably be better explained on the basis of an understanding of the syntactical characteristics of popular music than by prevalent theories of postmodernism; the latter appear problematic not only due to their speculative and unsubstantiated nature with regard to media reception processes (cf. Frith and Horne 1987, p. 11), but their explanatory value as regards syntactic features of music video also seems to be limited (cf. Frith 1988, p. 207).

In an earlier work, I have presented a theoretical discussion of the characteristics of popular music syntax and their consequences for the analysis of music video, as well as a general typology of relationships between music and visuals (see Björnberg 1992a). The purpose of the present paper is to modify and substantiate in more detail this previous account, by way of an analysis of structural relationships of music and images in a number of music videos. This analysis is aimed at demonstrating both the range of possible different types of such relationships, the limitations imposed upon them by musical syntax, and their relative 'openness' to varying kinds of reception and interpretation, depending on the mode of listening/viewing applied.

Postmodern society, postmodern media, postmodern audiences?

The specific characteristics of (the visual dimension of) music video attracting the attention of writers and scholars may be summarised as the breakdown of linear narrativity, of causal logic, and of temporal and spatial coherence. The apparently widely accepted mainstream line of explanation of these characteristic traits is based on the argument of postmodernism: the 'postmodern condition', pervading
all of contemporary industrialised society, finds its most adequate representation in the fragmentary forms of music video, combining practices from Classical high art, avant-garde modernism and popular culture (see, for instance, Aufderheide 1986; Kaplan 1987; Strøm 1989). The evaluation of the consequences of this situation varies among authors; whereas, for instance, Fiske (1986) regards the 'refusing of sense' as a practice potentially liberating from bourgeois hegemony of meaning, Tetzlaff (1986) is more pessimistic as to the existence of any such emancipatory potential. While postmodernism sometimes approaches the vague, all-embracing status of a Zeitgeist, many scholars relate it, and the characteristics of music video, to the development of the specific media codes of commercial television (Berland 1986; Jones 1988; Kaplan 1987; Larsen 1987), and especially, stressing the advertising function of music video, to the aesthetics of television advertising (Allan 1990; Frith 1988; Goodwin 1987; Kinder 1984; Laing 1985; Movin and Øberg 1990; Strøm 1989 and others). The loss, in this process, of rock music’s presupposed ‘authenticity’ of expression is also commented upon (Grossberg 1988); this is described both in negative terms (Movin and Øberg 1990), and approvingly, as offering the means for a ‘celebration of artifice’ (Ihlemann 1992).

The explanations for the emergence of music video in its specific form also include references to the significance of the development of technology, both television (cable) and video production/post-production technology, as well as that of musical production and reproduction (Berland 1986; Goodwin 1987; Ruud 1988). Ruud (ibid.) also discusses the connections with video art, emphasising the aesthetic quality of ambiguity inherent in the non-narrativity of music video (see also Strøm 1989).

Several explanations for non-narrativity from the point of view of the audience have also been offered. One line of argument apparently lying close at hand involves the view of the specificities of music video as a consequence of a general tendency, determined by social and media-technological developments, towards a Benjaminian ‘distracted mode of perception’, requiring a structure allowing the user to quickly catch the point of the message (Jones 1988; Larsen 1987; Movin and Øberg 1990); at times this view is extended into postmodernist speculation on the ‘schizophrenic subject’ (Jensen 1988). Another common explanation (presupposing the user’s devotion of a more continuous attention) emphasises the functions of fantasy, escape from everyday reality and imaginary problem solution provided by the incoherent and fragmentary structures of music video (Allan 1990; Aufderheide 1986; Morse 1986). The parallelisms between these structures and the workings of the primary processes of the psyche have been pointed out (Brown and Fiske 1987; Ruud 1988), as well as the potential of fantasy for exploitation for commercial purposes (Kinder 1984). In this context connections have also been made with the need-structures produced by the new, ‘narcissistic’ socialisation patterns suggested by German socialisation theory (Forsman 1986; Larsen 1987).

**Music: the missing connection**

Few of the theories related above can be totally dismissed as contributions to the understanding of the characteristics of music video. Nevertheless, in spite of the significant role of music in this context (if nothing else, as the alleged justification for the entire phenomenon), most of the authors cited have remarkably little to say about it. The opinion has been proposed that, for the user, the music is somehow
‘dominated’ by the visuals (Berland 1986; Kinder 1984; Tetzlaff 1986); this, however, seems to have little bearing on matters other than the individual mode of perception and musical preferences of these scholars. Others point out that historically, popular music performance has always been an audiovisual phenomenon, combining musical sound and visuals (Berland 1986; Goodwin 1987; Laing 1985), and several authors state that music is primary in relation to the visuals, although without elaborating much further on the nature and consequences of this primacy (Goodwin 1987; Laing 1985; Strøm 1989). Kaplan seems to regard music video visuals as mainly based on song lyrics (1987, pp. 47f.), restricting her discussion of musical characteristics to little more than one half page (1987, pp. 123f.).

In several instances writers on music video have given indications of the position regarding the conditions for and/or operation of musical signification underlying their respective treatment of the subject. Ruud (1988) regards signification in music as mainly determined by social context and conventionalised. Movin and Øberg, following Adorno, characterise music as non-referential and experienced by way of ‘emotional empathy’ (1990, p. 127), while Brown and Fiske state that ‘Rock videos, like rock music, work primarily on physical sensations and produce a physical pleasure that opposes the common sense of linear narrative in dominant ideology’ (1987, p. 62). Morse, in her discussion of music video functioning as ‘a form of magical thinking’; (1986, p. 24), touches upon aspects of music as performed (the star-singer ‘creating a visual world’ by virtue of his or her privileged position in the musical as well as visual mode of communication) but does not relate this to other musical dimensions. Whereas these various positions indicate relevant aspects, they are hardly exhaustive on the matter of musical signification in music video, and in most cases seem not to inspire the undertaking of more detailed analyses.

Allan, in a discussion of correspondences between musical film and music video, identifies the functions of music as ‘those of grounding, unity and resolution’ (1990, p. 9), while Kinder (1984) states that the continuous flow of the music imposes unity on the video. These are among the rare instances of music video scholars approaching the significance of structural aspects of music; another example is Berland’s (1986) somewhat impressionistic account of the relationship between formal song structure and visuals. Unfortunately, the embryo of structural analysis inherent in the latter approach is not followed up, probably due both to the author’s Adornian view of popular music structures and to the visual domination postulated (cf. above).

On a theoretical level, Frith’s (1988) discussion of structural characteristics of music video amounts to a well-considered counter-argument, based on musical facts, against postmodern theory speculation. Frith points to the general structuring principles of movement as ‘the metaphor for sound’ (1988, p. 216) and montage as representation of rock’s musical experiential qualities, and also discusses the relationships between musical and visual repetitivity. Although the arguments are coloured by the author’s obvious dislike for music video in its currently dominant form, and may be criticised on some points (cf. Björnberg 1992a, pp. 382, 386), this article remains an important corrective for writers on the subject. Frith’s arguments have also been picked up by Nielsen, stating that music video images are subordinated to the music by way of ‘a rhythmically determined pleasure principle’ (1991, p. 299), and Forsman, who adds to the montage/beat homology one between sound and an ‘associative spatiality’ (1991, p. 9).
However, despite these theoretical contributions, detailed analyses of music videos relating visuals to musical structure have been very sparsely presented. Ruud (1988) performs lengthy analyses of three videos; however, only one of these (Paul Simon's 'René and Georgette Magritte with their dog after the war') deals in greater detail with musical (mainly tonal) structure, thus emphasising unique individual features rather than general structural principles, and giving an unfortunate bias to the study. Larsen (1987), basing his discussion on Greimas's narrative theory, analyses the relationship between musical and visual syntax in Phil Collins' 'Against All Odds'. The author convincingly relates visual structure to musical segmentation and intensity (dynamics and texture) processes. The implied seeds of a more general application of narrative theory to popular music are, however, not followed up; in a later work (Larsen 1988), the author restricts his discussion to Western art music and its derivatives in Hollywood film music, by implication dismissing contemporary popular music as generally non-narrative and not susceptible to this kind of analysis.5

**Popular music form and narrativity**

In an attempt to remedy some of the shortcomings of the literature on music video recapitulated above, the analyses to be undertaken here will take as their theoretical point of departure the issue of the potential for signification in popular music syntax. That is, I am concerned with 'primary' rather than 'secondary' musical signification (Middleton 1990, pp. 220ff.), and with the signification of syntactical processes, as opposed to particular, individual musematic meaning (ibid., p. 235). I would also like to make a distinction in favour of the structural rather than the (in a psychoanalytical sense 'more primary', i.e. related to primary processes of the psyche) 'physical sensation', para-structural signification dimension of music; although, as Middleton (ibid., p. 219) points out, these cannot be clearly separated, the latter aspect of musical communication appears as yet to have attracted considerably more attention from writers on music video than the former.

Differently phrased, this is a matter of the potential for narrativity in music. In this connection, Middleton suggests a general distinction between three modes of construction of musical syntax:

In contrast to the narrative category, which privileges difference, there is what we can call an 'epic' mode, where the focus is on repetition and varied repetition; and in between comes a 'lyrical' category (marked by symmetrical open/closed and binary structures). (Middleton 1990, p. 216; my italics)

The positing of 'narrative' and 'epic' as polar opposites may appear somewhat confusing against the background of the everyday use of these terms in relation to literature and film; taking up the geometrical metaphors already hinted at by Middleton, a better terminology would perhaps be using 'linear' instead of 'narrative', 'circular' instead of 'epic', and maybe 'elliptical' (in the strictly geometrical sense) for 'lyrical', implying the dual aspects of movement and return inherent in the latter syntactical mode. Still, the categorisation appears valid and potentially fruitful for the analysis of popular music syntax, and although these terms in this usage may primarily be comprehended as rather abstract structural analogies, they nevertheless also arguably indicate experiential correlates to models of musical syntax, that is, to musical form, in a general sense. The formal
structures of popular music are usually conceived of and described in terms of standardised, 'neutral' structures, having achieved 'a naturalness' (Berland 1986, p. 44) to the listener; still, this does not render their intrinsic potential for signification non-existent.6

Middleton links the narrative mode of musical syntax construction with the absence of repetition, with the extreme variation of the 'infinite set', in which no musical element is repeated, the music continuously moving forward linearly and teleologically in a manner homologous to the structures of (verbal/visual) narrative. In order to investigate further the existence of homologies between music and narrative, structuralist narrative theory appears to offer some useful concepts.7 A basic distinction made within this body of theory is that between *story*, the signified or content of narrative, and *discourse*, the signifier or structure of the narrative text (Chatman 1978, p. 19; cf. Barthes 1977, p. 87; Genette 1980, p. 27). While the semantic precision of musical narrative as to its content, is 'story', is low compared to verbal or visual narration, musical structure may be demonstrated to be capable of exhibiting several of the characteristics constitutive of narrative discourse. In a somewhat loose but helpful phrasing, Chatman (1978, p. 25) identifies the distinctive features of narrative as 'eventhood, characterhood, and settinghood'. These concepts appear adequately applicable to musical discourse. A basic feature of Western music since the Renaissance, and particularly of popular music, is the structural dualism of melody and accompaniment (Maróthy 1974, p. 22; Tagg 1979, pp. 123f.). As Tagg indicates, this dualism is generally conceived in terms of a relationship between figure and background, or between individual/character and environment/setting. These are frequently consistently identifiable with separate musical 'voices'; however, the structural specificities and polysemic nature of music also allow for the same 'voice' to alternate between various functions, such as a guitar first playing a riff as part of the background, then switching to fill in the vocal line as a 'secondary character', and subsequently becoming the 'main character' in a guitar solo. As for the quality of 'eventhood', musical structure may appropriately be described as a succession (and/or simultaneity) of events, ranging from instantaneous changes to long-term processes. From this point of view, 'musical form' is defined by the particular nature of each musical event, by the temporal density of events and by their distribution throughout the duration of the piece in question.

For the purpose of investigating the *degree of narrativity* of particular musical structures a further distinction made by Chatman may be useful. Narrative discourse is characterised as being constructed from *statements*, which may be divided into *stasis* ('is') statements and *process* ('does') statements (Chatman 1978, pp. 31f.); a parallel distinction is the one made by Greimas between *radotage* and *affabulation* (cf. Larsen 1987, p. 89). Narrativity, in the common-sense meaning of the term, is linked to the quantity and density of process statements: a narrative where 'nothing happens' clearly has a 'non-narrative' character. Adequate musical analogies to this categorisation exist both in Ruwet's (1987, p. 16) differentiation between *parametric* and *non-parametric* elements in music, and in Middleton's (1990, p. 215) distinction between *binary/digital* and *analogue selection*. Parametric elements are characterised by being constant throughout a piece of music or switching between the two poles of an opposition, i.e. being subject to binary alternation, while non-parametric elements are characterised by 'a fairly large number of internal distinctions of the same dimensions', i.e. they tend towards analogue selection.
On the basis of the analogy between ‘process statement’ and ‘non-parametricness’, narrativity in music may thus be seen as related to the quantitative and qualitative significance of non-parametric organisation/analogue selection, in agreement with the argument of Middleton related above.

Although, as Ruwet points out, it is impossible to determine a priori whether a particular musical dimension has a parametric or non-parametric organisation, in Western music in general dimensions such as tempo, mode, instrumentation and timbre tend to be parametric, while melodic pitch, surface rhythm and harmonic tension are more or less non-parametric. Narrative qualities in Western music have thus primarily been associated with the dimensions of melody and harmony: melodic processes of aperture and closure combine with harmonic processes of tension and release, forming potentially long-ranging, forward-directed musical structures. Illustrative examples of this may be found in ‘extremely narrative’ music, such as, for instance, Wagner’s operas; in popular music, however, a high degree of such narrativity is seldom the case. Generally, popular music is characterised by strophic disposition, i.e. repetition of a small number of well-demarcated sections, and symmetrical construction, i.e. larger units being constructed by binary combination of smaller units (cf. Björnberg 1992b, p. 4). These characteristics imply a predominance of the lyrical mode: the use of symmetrical structures, whose well-roundedness works against directional linearity, is distinctive of this mode, and strophic disposition involves processes of repeated return to an experiential focal point or ‘point of departure’, processes whose reflective character contrasts both with the linearity of the narrative mode and with the short-term repetitiveness of the epic category. This lyricality is not, however, homogeneously prevalent across the popular music field; it is modified by different styles (and individual pieces of music) tending more or less towards linearity or circularity. It is also important in an analytical context to take into account the possibility for other dimensions than melody and harmony to produce narrativity in popular music, by virtue of being organised non-parametrically.

Analytical concepts and criteria

On the basis of considerations presented in the preceding discussion, in what follows I will propose a list of analytical dimensions relevant to the analysis of popular music form and its narrative characteristics. In the subsequent section, the concepts presented will be applied to the analysis of a number of music videos.

Discursive repetition/structure of lyrics/function

This dimension comprises most of the factors contributing to the constitution of musical form in a non-analytical, ‘intuitive’ sense. ‘Discursive repetition’ is a concept suggested by Middleton, and defined as ‘the repetition of larger units, at the level of the phrase, the sentence or even the complete section’ (Middleton 1990, p. 269). Finding the ‘main sections’ of a piece is largely a question of identifying the largest sections subject to (discursive) repetition. In principle, this could be done by using Ruwet’s (1987) paradigmatic method of analysis; in practice, this would entail a number of problems concerning the ‘equivalence’ of sections (how much variation is allowed for one section to be considered a ‘repetition’ of another?), and an intuitive classification in most cases yields the same results,
although much faster. Taking the temporal order of sections and the structure of lyrics into account, main and subordinate sections may then be classified according to their respective functions; these functions (cf. Björnberg 1992b, pp. 4f.) include verse (V), chorus (C), bridge (B), solo (S), break (Bre), introduction (In), interlude (II) and coda (Co).

Demarcation

A basic, albeit somewhat crude, measure of the distinctness of demarcation of the sections of a piece is the number of musical dimensions exhibiting change in the transition from one section to the next: the more dimensions that are affected by change of some kind, the more well-demarcated the section will be perceived to be. The musical dimensions in which changes may occur include lyrics (L), melody (M), harmony (H), vocal texture (V), accompaniment texture, (A), instrumentation (I), dynamics (D), tempo (Te) and tonal centre (To).

Symmetry

Symmetrical binary construction being the implicit norm of popular music, asymmetrical structures will be perceived precisely as deviations from a norm, that is, occurrences of asymmetry are perceptually marked as ‘events’. In general, the experiential effect of a deviation from symmetry is inversely related to its size; deviations ‘below bar-level’ not only disturb period and phrase structure but also regular metre (cf. Björnberg 1987, pp. 76f.). The categories of deviation from the symmetrical norm include prolongation, (P: addition of metrical units to a symmetrical structure), truncation, (T: subtraction of metrical units), and elision, (E: ‘overlapping’ of two symmetrical structures).

Musematic repetition

This concept is defined by Middleton (1990, p. 269) simply as ‘the repetition of musemes’; ‘museme’ is here to be understood in the sense used by Tagg (1979, p. 71) as ‘the basic unit of musical expression’, although, as pointed out by Middleton (1990, p. 189), ‘the nature and size of the museme need to be regarded flexibly’. In comparison to discursive repetition, musematic repetition is thus characterised by the repetition of smaller units, and in most cases also a greater number of repetitions; still, the necessary flexibility sometimes causes problems of delimitation between the two categories, problems which are best discussed in each particular context. According to the dimension affected, musematic repetition may be classified as melodic (M), harmonic (H), or accompanimental (A); however, in the interest of clarity of distinction it appears reasonable to exclude in this analysis those accompanimental parts which in popular music ordinarily feature musematic repetition to a considerable degree (i.e. drum, bass and chordal accompaniment parts; cf. ‘Motorial flow’ below).8

Directionality

This dimension involves the sense of forward-directedness effected by the evolving of melodic and harmonic structures through time (a more complete designation
would be pitch-related directionality). The proposal of this concept constitutes an attempt to summarise the effects both of parametric dimensions such as mode ('tonal language') and of non-parametric aspects of tonal organisation. It is thus related to the amplitude and frequency of fluctuations of melodic/harmonic tension, and partly analogous to the concept of 'prolongation' proposed by Lerdahl & Jackendoff (1984, pp. 242ff.). While directionality in principle is a quantifiable dimension, a detailed quantification would seem to entail a high degree of subjective judgement. I will therefore in my analyses only estimate the degree of directionality on a coarse quantitative scale, a high degree being considered as related both to the total number of different elements (melodic pitches and chords) used, and to the degree of adherence to the voice-leading principles of functional tonality (melodic leading notes, melodic dissonance/consonance progressions, 'strong' harmonic progressions, e.g. involving descending-fifth root progressions, 'regular' use of chord inversions, etc.). The degree of directionality is also affected by deviations from metrical symmetry, e.g. prolongation increasing the directionality of a dominant chord) and by musematic repetition (a 'strong' progression, when repeated, gradually loses its sense of forward-directedness). In connection with the analysis of directionality, occurrences of closure (cadential effects), classifiable as melodic (M) and/or harmonic (H), may also be conveniently identified.

Motorial flow

The presence of an accompanimental 'motorial continuum', mainly effected by drums, bass and chordal instruments, approaches the status of yet another popular music norm (cf. Brolinson & Larsen 1981, p. 200). Variations in the motorial flow are therefore also perceived as musical 'events'. As is the case with directionality, the degree of motorial flow is also quantifiable in principle but open to subjective judgement; the rough quantitative estimate in my analyses is based, on the one hand, on the number of contributing layers of musical texture, and, on the other, on the temporal density and regularity of the sound-events (beats) constituting the motorial continuum. It should perhaps be pointed out here that such an analysis presupposes a mode of listening involving the expectation of an explicit 'spelling out' of motorial flow; such a mode of listening may arguably be assumed to be frequently applied to contemporary popular music, in contrast with, for instance, a jazz-oriented mode of listening enabling the experience of intense motorial flow as implicit in a single melodic line (cf. Durgnat 1971, pp. 36f.).

Dynamics

Changes in the overall dynamic level of the music are estimated on a coarse quantitative scale.

Sound processes

This dimension involves changes affecting the characteristics of the overall sound or of the sound of individual voices/instruments, e.g. changes in timbral quality, amount of reverberation or echo, dynamic balance between individual voices/instruments, etc.
Table 1

<table>
<thead>
<tr>
<th>Discursive repetition</th>
<th>Narrative/ linear</th>
<th>Lyrical/ elliptical</th>
<th>Epic/ circular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demarcation</td>
<td>no</td>
<td>yes</td>
<td>Some</td>
</tr>
<tr>
<td>Symmetry</td>
<td>variable</td>
<td>distinct</td>
<td>indistinct</td>
</tr>
<tr>
<td>Musematic repetition</td>
<td>low</td>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>Directionality</td>
<td>no</td>
<td>some</td>
<td>yes</td>
</tr>
<tr>
<td>Motorial flow</td>
<td>high</td>
<td>variable</td>
<td>low</td>
</tr>
<tr>
<td>Dynamics</td>
<td>variable</td>
<td>relatively constant</td>
<td>constant</td>
</tr>
<tr>
<td>Sound processes</td>
<td>short-term</td>
<td>long-term</td>
<td>constant</td>
</tr>
<tr>
<td>IPF</td>
<td>high</td>
<td>relatively high</td>
<td>low</td>
</tr>
</tbody>
</table>

**Individuality predominance factor (IPF)**

This is a rough quantitative measure of the significance of ‘individuality’ in a musical piece, based on a reading of the dualism of lead vocal melody and instrumental accompaniment in terms of the individual/background relationship (cf. above; also see Björnberg 1987, p. 219). It is calculated in three ways: a) as the ratio of the total duration of vocal phrases to the total duration of sections featuring vocal parts, i.e. the duration of vocal phrases plus the rests between them (IPFv,); b) as the ratio of the total duration of sections featuring vocal parts to the total duration of the piece (IPFa); c) as the ratio of the total duration of vocal phrases to the total duration of the piece (IPFv, equal to IPFv, IPFa).

The general relationships between these analytical dimensions and the three modes of musical syntactical construction discussed above are shown in Table 1. The table entries indicate how the analytical dimensions, each one viewed separately, are related to the syntactical modes; however, due to the multidimensionality of music, actual pieces of music offer several possibilities for tendencies in different dimensions to work in different directions. A thorough theoretical analysis of the overall effects of such conflicting tendencies being a very complicated task, in the present context I will restrict myself to a discussion of such conflicts appearing in the particular examples to be analysed.

For the sake of clarity of distinction, in what follows I will partially effect the change of terminology suggested above, using the terms ‘linear’, ‘elliptical’ and ‘circular’ when exclusively discussing musical syntax, while retaining the narrative/lyrical/epic triad when speaking of visuals, combinations of music and visuals, or experiential aspects of musical structures. An initial hypothesis underlying my analyses is that the visual structures of music video are generally determined by, and reflect, the elliptical (i.e. non-linear) nature of popular music syntax, and the polysemic nature, or ‘conditional referentiality’ (Björnberg 1992a, p. 380), of musical denotational signification, rather than a ‘postmodern fragmentation’ of visual narrative. However, in light of the range of variation per genera et species characterising the syntactical construction of actual popular music, the conclusion that ‘rock music cannot in principle be illustrated with logical narrative processes’ (Nielsen 1991, p. 298; my italics) appears too rash; furthermore, visualisation in music video is of course not necessarily entirely determined by musical syntax. In the reciprocal interplay between musical and visual signifying systems there is
scope for various kinds of relationships between the two (or three, including verbal lyrics), and one of the purposes of the analysis is precisely the investigation of the practical and aesthetic limitations on this scope.

The objects chosen for analysis are four recent music videos recorded for MTV Europe. These have not been selected with an aspiration to any kind of statistical representativity, but with the aim of demonstrating a range of possible relationships between musical and visual structures. The musical analyses are also presented graphically (see Appendix, Figures 1–4).

A streetcar named Defier

The first object selected for analysis is the video for Bruce Springsteen’s ‘Human Touch’. The song is a ‘trad rock’ number in medium tempo and ‘VCB-form’ (VVCVCBVCC; see Figure 1); however, this is interspersed with several instrumental sections (In, II and S sections), giving the song an overall length of roughly 6'25". The last C section (preceding the extended instrumental Co) is prolonged by way of varied repetition of the last phrase. Demarcation is generally distinct, with changes in vocal texture, instrumentation and accompaniment texture supporting the segmentation defined by lyrics, melody and harmony. Sections are mostly symmetrically constructed, but two-bar prolongations occur in the C, B and S2 sections. Melodic musematic repetition is not very prominent; the V section melody may be characterised as using ‘varied musematic repetition’, with a small number of repetitions. The harmonic structure, though, utilises musematic repetition for the most part of the song: the V and C sections are based on the two-bar Mixolydian chord sequences | I | bVII IV | and | bVII | IV |, respectively, the first of which is also used in the In, II and S2 sections, while the second appears in the S1 section. The vocal sections, however, conclude with the sequence | bVII | bVII IV I |, producing a (weak) harmonic closure, combined with melodic closure in C sections. Directionality is thus fairly low throughout, except for the B section, featuring a more active (also Mixolydian) chord sequence; also, there is no final closure, since the song ends with a fadeout. Motorial flow and dynamics are highly variable, with variations reinforcing segmentation. Sound processes mainly concern the relative dynamic levels of lead vocals and lead guitar. The overall IPF stated is fairly low, due to the extended instrumental sections; however, an experientially relevant value would rather be higher, on account of these sections featuring improvisatory vocal elements.

All in all, the song exhibits structural features typical of the ‘epic’ (this time, in the sense ‘heroic narrative’) trad rock style characteristic of Springsteen. The basic conception is circular (the harmonic ostinatos and mainly low-directional melody), but this is modified by its adaptation to an elliptical framework (clearly perceptible, albeit weak, harmonic and melodic closures, discursive repetition of well-demarcated sections). The circular character, however, is also manifest in the relationship between the sections here termed ‘verse’ and ‘chorus’: the musical material of the C section mainly amounts to permuted V material (my identification of this section as a ‘chorus’ is based on temporal order of presentation, closure, and lyrics, the song’s title appearing at the end of this section’ cf. Björnberg 1992b, pp. 2f.). Thus there is little of the contrast or directionality usually associated with verse-chorus progressions; the largest contrast in the piece with regard to tonal processes is produced by the B section. The returns (discursive repetition) of entire
sections, being generally typical of the elliptical mode of musical syntax construction, have the effect of ‘grounding’ the processual flow by way of the ‘reconciliation of difference’ function discussed by Middleton (1990, pp. 223f.). Nevertheless, the song also has a distinctly narrative quality, effected by means of a cumulative succession of multiple circular/elliptical sections modulated by processes of variable motorial flow and dynamics (especially evident in the transition from II4 and II5, the most climactic point in the song), but this narrativity is thus of a particular, static, long-term, slowly evolving nature.11

The visuals of the ‘Human Touch’ video, as is the case with most music videos, are to a large extent structured by the segmentation of the music; hence the following account is arranged according to corresponding musical sections:

In: Nocturnal North American city street; approaching streetcar.

V1: Superimposition of image of Bruce Springsteen (hereafter, Springsteen1) lip-synching vocals, shower of sparks from trolley-wire on the word ‘pretenders’; dissolve to streetcar, with man in window easily identifiable as Springsteen (Springsteen2).

V2: Intercutting between Springsteen1 and passengers of streetcar, including Springsteen2.

C1: Springsteen1 sings against indeterminable grey-brownish background; dissolve to streetcar running; sparks on the word ‘touch’.

II1: Rapid montage composed of shots of running policemen, guitar-player (Springsteen3) in tiled pedestrian tunnel or suchlike, Springsteen2 in streetcar and young people in nocturnal city street settings illuminated by flashing lights.

V3: As V2; rapid insert of indistinct close-up of woman on the word ‘tonight’.

C2: Springsteen1 singing; lonely woman in room; Springsteen2 in streetcar exchanging smiles with little girl.

II2: Rapid montage of Springsteen3 and street scenes; restless camera; flashing lights.

B: Montage of Springsteen3 lip-synching and street scenes; trolley-wire sparks on the word ‘we’re’ (‘. . . . all riders on this train’).

S1: Montage of Springsteen3 and street scenes, ecclesiastical scenes (funeral, christening), mother with baby, little girls dancing and wedding-couple.

II3: Montage continued; cut to tilted-camera shot of shadow of streetcar moving on housefronts; cut to Springsteen3.

V4, C3, C4, II4: Intercutting between Springsteen1 lip-synching and streetcar stopping, Springsteen2 getting off, crossing street, entering building, climbing stairs, walking down corridor, knocking on door.

II5: Rapid intercutting: Springsteen3 shouts; woman runs towards door, bare-chested Springsteen3 swings guitar; Springsteen2 enters room; Springsteen3; Springsteen2 and woman embrace; Springsteen1, now playing guitar, etc.

II6: Springsteen1 playing guitar; montage of this, caresses of naked body, and Springsteen3 with hands lifted as if in prayer; cut to trolley-wire sparks with Springsteen3 in background on percussion intro to next section.

S2: Montage continued, mostly featuring caresses and Springsteen3; however, the video ends with Springsteen1 smiling relaxedly and self-confidently into the camera.

As indicated by this recapitulation, the video is composed of three distinct visual components: an act of narration, a visual narrative, and a series of disconnected scenes in various settings, i.e. in agreement with Kinder’s (1984) well-
known categorisation. Each of these three components features its own edition of Bruce Springsteen: Springsteen1, the balladeer and narrator; Springsteen2, the actor and protagonist of the narrative; and Springsteen3, the 'guitar hero' from the streets. The first two components together make up the visuals during the In section, most of the vocal sections and the Il4 section (which may be regarded as a short 'coda' to the prolonged C4 section), while the third dominates the instrumental sections III–3 and S1, as well as the vocal B section. In the montage of the extended coda section (Il5/I6/S2), the three components seem to merge and form a synthesis. Accordingly, the separate identities of the three Springsteens: the calm, controlled and controlling narrator, the vulnerable protagonist travelling to see his loved one, and the defiant street rebel-musician, apparently move towards becoming different aspects of one and the same identity (the variations of the clothing of Springsteens 1 and 3 are interesting in this respect), although the concluding shot of the video indicates which of the three is to be perceived as the 'real' one (cf. Frith 1988, p. 216).  

In accordance with the general character of the musical narrativity of the piece, the visual narrative is fairly low on process statements, i.e. not very much 'happens': a man travels by streetcar to meet a woman waiting, they meet and rejoice. Nevertheless, it forms a coherent 'logical narrative process', at the same time as its openness and ambiguity (who are this man and this woman? where does he come from? for how long have they been separated? will they have to part again?) allows for a range of interpretations (cf. Goodwin 1987, pp. 26f.). In the musical sections high in directionality and/or motorial flow the visuals evolve into an 'elliptical' (this time, in the literary sense) narrative, inviting paraphrases in the direction of 'it's a dangerous world out there, but human touch is still essential in all stages of life'. Thus a conflict is set up, musically as well as lyrically/visually, but this conflict is reconciled, visually/narratively as well as musically, in the coda of the song.  

**Blue dolphin beat**

My second example is the video for Bryan Adams' 'Thought I'd Died and Gone to Heaven'. This song is a mainstream rock ballad in 'VCB-form' (VCVCBVCC; see Figure 2), where the vocals of the first eight measures of the V3 section are replaced by a guitar solo. Demarcation is distinct: the segmentation defined by lyrics, melody and harmony is reinforced by changes in vocal texture, instrumentation and accompaniment texture. Deviations from symmetry are numerous in the form of 2½-bar prolongations of V sections (two bars by phrase repetition, one half 'extra' measure added) and one-bar prolongations of C sections (1½ bar in C3). Musematic repetition is sparsely used; the C sections feature the two-bar, four-chord vamp | I V | vi IV |, which may be regarded as a 'border case' of musematic repetition (cf. Björnberg 1992b, p. 12), while the $V_b$ sections contain a discreet guitar riff. Experientially, the combination of In and the first eight bars of V1 also produce an effect of harmonic musematic repetition, an effect not present in the following V sections. Directionality is medium-high and fairly variable, with lows in $V_x$ sections and a peak in the B section. No unambiguous melodic or harmonic closure occurs, and the song ends with a fadeout. Motorial flow and dynamics are nearly constant throughout, except for breaks at the ends of sections and the first four bars of the B section. Sound processes include a temporary echo effect on the first entry of the lead vocals, and a successive domination of
accompaniment over lead vocals in the concluding C sections. The overall IPF is fairly low, in spite of vocal sections constituting the bulk of the piece, on account of the ‘gaps’ in the vocal line (IPF_v = .63).

As is the case with most rock ballads, the song is predominantly elliptically constructed, with an orderly repetition of clearly demarcated sections. This, however, is imprinted with distinctly linear traits: the fairly high and variable directionality, and the frequent deviations from symmetry,\textsuperscript{13} while motorial flow, dynamics and sound processes support ellipticality. On the whole, the processual character of the song is thus mainly lyrical, i.e. reflective and self-contained, but with a not insignificant narrative potential, although this narrativity is primarily of a ‘local’, short-range nature, and accommodated within the larger elliptical framework.

The visuals of the video, arranged according to musical segmentation, are summarised in the following:

\textit{In:} One continuous shot; camera, directed downwards, moves over grassy meadow at night but with blue-ish lighting (which pervades the entire video); camera tilts up to horizontal, dollying in on Adams, standing in the meadow.

\textit{V1:} Shot continues; camera stops at m. 14; three dolphins jump out of the grass and fall back, in slow motion; camera dollies upwards and out; cut (for the first time, at m. 19) to medium close-up of Adams; cut to drummer, then Adams, camera dollying down/in, then out; cut to Adams from another angle, camera moving in; cut to close-up of Adams.

\textit{C1:} Dissolve to distance shot of entire group, upwards down; camera dollies in while turning clockwise; dissolve to Adams, camera moving down/out, showing guitarist; wind moves the previously calm grass; at m. 33 camera stops and marathon runners (?) start appearing out of the grass at gradually decreasing distance; the last six or seven of these, who resemble the dolphins by wearing sweaters pulled up over their heads, disappear at top of picture; camera continues to move out; cut to drummer, camera moving upwards and out while turning.

\textit{V2:} Similar images of Adams and the other musicians, camera constantly in movement; at m. 47 a dolphin jumps up, falling down two shots later; last shot in section featuring bass player, with camera moving in, then around, then upwards and out.

\textit{C2:} Similar images; section ends with fade to black.

\textit{B:} Fade in to medium close-up of Adams superimposed on meadow with moon and clouds above; the musicians, superimposed, act swimming- (or stage-?) divers, fully dressed, at top of picture; from m. 69 montage as before, with flying birds superimposed from m. 73.

\textit{S:} Dissolve to camera dollying in, through shrubbery, to guitarist with moon above; camera stops at m. 81, moon turns into strobe light and a dolphin jumps up and somersaults, landing with great splash of water; cut to close-up of Adams.

\textit{V3:} Montage as before; in last shot of section, one dolphin jumps up.

\textit{C3:} Similar montage; in second shot of section, three dolphins come down.

\textit{C4:} Similar montage.

\textit{Co(C5):} One continuous shot; dissolve to medium close-up of Adams; camera moves out, showing the other musicians; at m. 121 five dolphins jump out of the grass in front of group, the middle one, corresponding to Adams’ position in the group, jumping later and higher than the others; at m. 123, the dolphins
having disappeared, the camera sinks, still horizontally directed, down into the grass.

Obviously, this is not a ‘narrative’ in any qualified sense. The main visual components remain the same throughout the video: the nocturnal setting in the meadow, the blue-ish illumination, and the musicians doing nothing but performing (or, rather, lip- and hand-synching the performance of) the song. The things actually ‘happening’, i.e. the appearance and movements of the dolphins, runners and birds inserted by means of technical special effects, stand out as isolated occurrences seemingly impossible to relate to the other visual components (or to the song’s lyrics) within any coherent and logical narrative framework. One ‘narrative’ element in the video, however, is constituted by the almost constantly moving camera, producing perpetual shifts in distance and orientation; still, the ‘story told’ by these camera movements remains completely static.

Of course, the video is not intended to present any independent narrative, but to be experienced in connection with the music of the song. In spite of my arranging the summary above by musical segmentation, the visuals of this video are not very clearly structured by this segmentation, nor by the motorial layers of the music. Visual transitions are produced by dissolves rather than straight cuts and often displaced in relation to the beat (which is continuously spelt out by the bass, playing straight quavers, and drums). Some transitions (C1-V2, V2-C2; cf. above) are marked by particularly energetic camera movements, but generally the visual processes (besides the camera work, also the slow-motion movements of dolphins and runners) emphasise continuity and constancy, in a ‘relay’ relationship with musical processes (cf. Barthes 1977, pp. 38ff.). In this way, the visuals underline the lyrical, reflective character of the elliptical musical structure, potentially opening up vast fields of association to the user. On the other hand, except for the work of the ‘narrative camera’, the narrative potential of the music is not realised in the video; this, however, is not so much a matter of the images ‘dominating’ the music as of the visuals focusing attention on some aspects of the music while concealing others.

**Egyptian divorce**

The object for my next analysis is the video for Michael Jackson’s ‘Remember the Time’. The song is a typical Jacksonian hip-hop-influenced dance piece in ‘VCB-form’ (VCVCBCCCCC; see Figure 3). Furthermore, there is a soloistic/instrumental middle segment (the Bre and II sections; m. 73–104); these are peculiar to the video version, not being present in the CD album version of the song. Demarcation is fairly indistinct, with mostly only changes in instrumentation (in most cases rather subtle) and vocal texture supporting the segmentation defined by lyrics and melody. Sections are symmetrically constructed throughout the piece, exhibiting no deviations. Harmonic musematic repetition is prominent; except for the B section and the (implied) static harmony of the Bre section, the entire song is built on a four-bar V-i ostinato (with varying alterations of the V chord). Although the vocal line shows general repetitive tendencies, actual melodic musematic repetition (more or less varied) occurs only in the Bre, II and concluding C sections. In spite of the strong V-i progressions, directionality is low, due to the large number of repetitions, and no melodic or harmonic closure occurs, the song ending with an echo effect on an altered V chord. Motorial flow and dynamics exhibit little variation, with the exception of the Bre section. Sound processes
mainly affect presence and reverberation of the vocals. The IPF is high, due both
to rests between vocal phrases being short and to most sections featuring vocal
melody (Jackson’s habit of singing all vocal parts himself renders a clear distinction
between ‘lead’ and ‘background’ vocals problematic; the experiential effect is one
of ‘multiple lead vocals’).

The song’s structure takes a mainly circular form: the repeated harmonic riff
and constant motorial flow, acting against any sense of processual development,
serve to level out contrasts between sections; the ‘verse’/‘chorus’ distinction there-
fore mainly depends solely on lyrics and, to a limited extent, melody. This renders
the processual character of the song epic, i.e. static and ‘mythic’. The Bre and Il
sections provide some potential for the kind of long-range narrativity discussed in
connection with ‘Human Touch’, but to a considerably lower degree than in that
case.

As in many other Michael Jackson videos, the length of the ‘Remember the
Time’ video significantly exceeds that of the song, due to the presence of a
visual/dramatic intro and a (shorter) visual coda; these are denoted ‘Vis. In’ and
‘Vis. Co’, respectively, in the following summary:

Vis. In: Throne-room of Pharaoh in ancient Egypt, Pharaoh and Queen seated on
thrones; hooded figure approaches; Pharaoh asks ‘Now, what is it you’re going
to do?’ (the only instance of spoken dialogue); hooded figure takes brownish
dust from pouch and spreads it on the floor, where it moves into circular pattern;
hooded figure steps onto pattern and disappears mysteriously through the floor,
leaving only his cloak; where he’s stood, an amorphous shining figure arises,
eventually metamorphosing into Michael Jackson. The scene is accompanied by
Phrygian ‘suspense’ underscore music played by strings and electric guitar, and
by sound effects.

In: Shots of Jackson, moving rhythmically to the beat, alternating with coun-
tershots of Pharaoh and Queen.

V1: Jackson lip-synching vocals; countershots of Queen; Jackson kisses her hand;
Pharaoh signals to his guards; Jackson turns and runs down the stairs from the
thrones, leaving Queen with hand outstretched.

C1: Intercutting of guards running or pointing, Pharaoh, Queen swooning; at end
of section cut to doorway where Jackson appears, swirling on percussion pick-up
to next section.

V2: Intercutting of guards searching, Jackson with passing camel in front, Jackson
in harem, snakes in basket.

C2: Jackson dancing in circle of women; at end of section continuity cut from veil
of dancing women to drapery.

B: Pan from drapery to Queen on bed; zoom out reveals Jackson at other end of
room; intercutting of Queen and Jackson.

C3: Queen on balcony, looking out onto pyramids; Jackson approaches; they
embrace and kiss; at end of section cut to column, behind which arm protrudes
making ‘Egyptian’ gesture.

Bre: Intercutting of guards and dancers emerging from behind columns; intercut-
ting of Jackson and dancers; zoom out from Jackson revealing surrounding
dancers.

Il: Dance scene; Jackson and dancers from various angles.

C4, C5, C6, C7: Dance scene continued; at end of last section cut to close-up of
Jackson.

Co: Jackson swirls; cut during swirl to different background.
Vis. Co: (accompanied by percussion underscore in same tempo as song); Jackson stops and looks around, finding himself back in throne room; Pharaoh in from left; Jackson smiles impudently and makes a dash; menacing-looking guards approach; Jackson, finding himself surrounded, escapes by swirling and disappearing mysteriously.

In this case, there is no equivocality as to the narrativity of the visuals: a logically consistent narrative, with a considerable degree of temporal and spatial coherence, is presented by means of well-established filmic codes. Nevertheless, this narrative is clearly divided into segments exhibiting varying levels of 'realism'. In the first, and 'most realistic' segment, the narrative evolves in 'real time'; this segment constitutes the 'visual intro'. In the second segment, corresponding to musical sections In-C3, the realism is modified by a more ambiguous construction of temporal continuity. The third segment, starting at the beginning of the Bre section, dismisses realism completely, suspending the narrative flow throughout the remainder of the song; the second shot of the guards in that section constitutes the last 'realistic' element. The 'visual coda', by returning to the first level of realism, resumes and concludes the narrative.

This segmentation of the visuals is obviously related to musical segmentation. In effect, this video constitutes a condensed version of the classical Hollywood musical film (as seems to be a favourite procedure with Michael Jackson), underlining the parallelisms between the two forms discussed by Allan (1990), and lying significantly closer to the Hollywood musical than Morse (1986, p. 23) considers typical of music video. The correlation, characteristic of the musical film, of 'realistic narrative' with absence of music and of 'fantasy' with its presence is, however, modified in the video by the establishment of two levels of realism in the former mode, one without music (except for the underscore) and one with music. The 'fantasy' in the video (the dance scene), the transitions in and out of which are musically marked by echo effects (in the Bre and Co sections), corresponds in function to the musical number of the musical film. However, the connection of 'level 2' narrative with circular music structure means that the latter is not very clearly realised in the visuals; with a few exceptions, edits and depicted movements are not strongly co-ordinated with the music. Due to the 'narrative interest' of the images, there is thus, in this particular segment of this particular video, some reason to speak of 'domination' of the music by the visuals. The only 'narrative' musical element present here is the high IPF, enabling in a limited way a musical support for the evolution of visual narrative.

A space obscurity

My final analysis concerns the video for Snap's 'Rhythm is a Dancer'. This song is a pop-techno dance number; the overall formal structure might be characterised as an instance of 'expanded standard form' (cf. Björnberg 1992b, p. 2) with the disposition VVBV (see Figure 4), although the construct is not supported by the tonal processes typical of that form-type. The S (or B) section is composed of eight bars of instrumental solo followed by a sixteen-bar rap. Demarcation is of medium distinctness, the segmentation defined by lyrics and melody being supported by changes in instrumentation and accompaniment texture. All sections are symmetrically constructed with no deviations. Musematic repetition predominates in harmony and accompaniment structure: the bulk of the piece uses the two-bar Aeolian chord ostinato | i bVI | bVII i |, spelt out by a synthesiser riff in quavers. The
vocal melody, although exhibiting no clear musematic repetition, works repetitive tendencies into arch-shaped lines. Directionality is generally insignificant, with a low in the Phrygian-coloured static-harmony B section. The final tonic minor chord gives no strong sense of harmonic closure, since it appears in the same metric position as previously in the harmonic riff; the song ends abruptly with a percussion echo effect disrupting regular metre by a shift to a rhythm of dotted quavers. Motorial flow is built up in discrete steps up to m. 32, thereafter being mostly constant, with the exception of the B sections. Dynamics show little variation; sound processes primarily concern the amount and quality of reverberation applied to vocals and/or overall sound. The IPF is medium-high as regards both rests between vocal phrases and relative duration of vocal sections.

Thus, like the preceding example, this song too exhibits a circular structure based on repetitive harmony and fairly constant motorial flow. One important difference, though, is that the application of the 'verse/chorus principle' is even less distinctly manifest in 'Rhythm is a Dancer'. While the vocal melody effects a demarcation in the middle of the V section, implying a differentiation into a 'verse' \((V_A)\) and a 'refrain' \((V_B)\), this is not supported by any variation as regards lyrics between the different \(V_A\) sections (all lyrics remain identical in all three \(V\) sections). The epic character of the song is therefore somewhat more emphasised in comparison with 'Remember the Time'; however, by the alternation of sections and the variations of motorial flow, the song still provides some (small) potential for the previously discussed long-range type of narrativity.

The visuals of the video, arranged as before according to musical segmentation, are summarised below:

**In:** Establishing shots and show pans of rocket launching site, illuminated by pale mauve-orange lighting and occasional white flashes and spotlights; rockets standing in background; smoke; superimposed sped-up moving clouds; female singer, in tight, shiny, black dress, on slowly-rising construction-elevator; surrounding people dressed in similar garments of dull metallic lustre.

**V1:** Similar scenery; three persons holding huge spheroids over their heads; mechanical body movements; bass player on similar slowly-rising construction-elevator behind and to the left of the one holding the singer.

**V2:** Cuts on downbeat of every second bar; on m. 33 to computer/paintbox graphics with symmetric mirror effect; on m. 35 to different graphics featuring moving clouds, human silhouette in slow motion and CAD (computer-aided design) rotating wheel; on m. 37 to singer; on m. 39 to graphics resembling the first but with two human silhouettes and moving 'atom' symbol; then to intercuttings of singer, lip-synching and gradually moving more intensely, and bass player.

**Bre1:** Scenery similar to that of the V1 section.

**S:** New graphics: superimpositions of human silhouettes running in slow motion, clouds and nautical charts, changing by quick dissolves in quaver rhythm; on M. 61 cut to graphics featuring zoom in on human silhouette in slow motion and CAD rotating sphere; from m. 65 montage of these elements, medium-close-ups of rapper lip-synching and humans moving mechanically; at end of section 'zoom in cut' on face of rapper on the word 'dancer'.

**V3:** Scenery similar to that of section V1; several slow pans.

**Il:** Cut to human silhouette/nautical charts graphics; on m. 101 cut to bass player.

**Bre2:** Cuts on downbeat of every second bar to various graphics; on m. 112 cut to CAD sphere with superimposed clouds.

**Co:** Sphere recedes into distance, rotating; fade to black.
As in the case of 'Thought I'd Died and Gone to Heaven', this video is manifestly non-narrative. Two main visual components are set up: on the one hand the launching site scenery, on the other hand various types of special-effect graphics, but none of these shows any significant tendency towards narrative evolution.

Thus, to an even higher degree than in the Bryan Adams example, the visuals of this video are dependent on the music. The movements depicted and the editing both work towards an emulation of musical rhythm, the superimposition of several different types of movement reflecting the latter's multi-layered nature. Many of the cuts described above as 'on the downbeat' are actually placed fractionally before the beat, thus producing an off-beat effect entering into a relay relationship with, and reinforcing, the rhythmic flow of the music. No narrative components interfere with the alignment of the visuals with the epic character of the piece. The actual content of the visuals potentially opens up large fields of association, but these are not strongly structured by the visuals nor by the music of the video.

Towards a typology of structural relationships

On the basis of the preceding analyses, I would like to propose a tentative typology of structural relationships between music and visuals in music video (see Table 2). The four types indicated in the table are, however, not equally representative of contemporary music videos. As Middleton (1990, p. 217) points out, post-rock'n'roll popular music has generally tended towards mixtures of the elliptical and circular modes of syntax construction. The music/visuals relationships of most music videos could therefore be expected to approach one of the two types in the right-hand column of the table. However, this typology, being rather schematic, cannot be expected to exhaust the range of possible relationships, and it evidently needs further modification and differentiation. In this context, it should also be noted that in relation to the videos analysed here, the concepts 'narrative music' and 'epic music' must be regarded on a relative scale, as each of the four songs is based on a clearly discernible elliptical structure; on the whole, decidedly linear or circular structures rarely appear in popular music (cf. above). Thus, it appears particularly pertinent for the further elaboration of this typology to focus on the elliptical mode of musical syntax construction, to which no obvious visual counterpart seems to exist. The musical function of providing 'unity', mentioned by Allan and Kinder (cf. above) is relevant in this context, but conceived in too general terms: since music always has a tendency to effect unity by means of sheer redundancy in comparison to other communicative modes (cf. Björnberg 1992a, pp. 382f.), the crucial point is rather how this unification is brought about in each particular case.

It appears, however, to make some sense to draw a parallel between the linear/elliptical/circular axis and Kinder's (1984) aforementioned 'narrative/performance/dreamlike visuals' axis. The alignment of 'linear' with 'narrative', and of 'circular' with 'dreamlike visuals', seems relevant, as indicated by my analyses. Furthermore, performance visuals in music video often appear 'neutral' with relation to the signification of musical syntax, and they may thus constitute a 'natural' or privileged mode of underlining the reflective, introspective, reconciliation-of-difference character of the elliptical mode of musical syntax construction. These musical characteristics, described by Moore (1993, p. 85) in terms of the
Table 2

<table>
<thead>
<tr>
<th>VISUALS</th>
<th>MUSIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narrative</td>
<td>'Film music' relationship, i.e. narrative processes in music and visuals interacting, but on terms set by the music rather than the visuals</td>
</tr>
<tr>
<td>Epic</td>
<td>Assignment of 'mood music' function to music, i.e. the music providing backcloth to the visual narrative</td>
</tr>
<tr>
<td>Narrative</td>
<td>'Circularisation' of music, i.e. focusing of attention on non-narrative aspects of the music, concealing its narrative potentials</td>
</tr>
<tr>
<td>Epic</td>
<td>'Musicalisation' of visuals, i.e. the visuals primarily reinforcing musical experience by synaesthetic 'translation' and emulation of musical experiential qualities</td>
</tr>
</tbody>
</table>

'fictionality' of the popular song as a 'self-contained entity', could thus be viewed as homologous to musical performance in its function as an 'irrational' aesthetic activity, detached from everyday reality.

The complexity of the issue should also be acknowledged by taking verbal lyrics into account. I have not commented much on the lyrics of the songs analysed here, since these seem to me to be more important to an analysis of content than to one of structure; still, the matter clearly calls for further attention. In this connection, it can be noted that music videos with a visual narrative are often based on narrative lyrics (or lyrics with some narrative 'seed' in them), to a large extent regardless of the degree of musical narrativity (cf. videos such as Rolling Stones' 'Under Cover of the Night', 'Small Town Boy' by Bronski Beat, and Tom Petty's 'Into the Great Wide Open'). Nevertheless, clear correlations as regards narrativity in different signifying systems appear to exist at least in some musical genres (notably the predilection for 'mythically' conceived narrative lyrics, narrative videos and 'baroque' musical syntax in heavy metal music; cf. Walser 1992).

Conclusion

None of the music videos analysed in this article is particularly exceptional in respect of aesthetics, neither as regards music nor visuals; rather, to my view, all four examples are (musically as well as visually) well-crafted, and quite effective, instances of mainstream forms of expression. The widespread reluctance to deal with such 'naturalised' mainstream forms, implicitly dismissing them as uninteresting and aesthetically deficient, seems to be partly rooted in a somewhat romantic belief in the subversive power of rock music (see, for instance, Berland 1986, Tetzlaff 1986). However, as Goodwin (1987, p. 31) indicates, the countercultural potential of pop culture is often exaggerated (especially, it seems to me, by US-American scholars and writers); furthermore, a full understanding of the subversive and the exceptional cannot be expected to be gained without also examining the 'ordinary' (cf. Tagg 1982, pp. 63f.).
To be fair, the confusions and theoretical disagreements discussed here concerning the characteristics of music video are partially dependent on the time of writing, as the dominant forms of music video have changed considerably over the years. Furthermore, music video being a very heterogeneous phenomenon, it sometimes also appears that ‘anything can be proven’, depending on how examples are chosen. This illustrates the extent to which the field of contemporary popular music is stylistically, socially and semiotically fragmented; still, despite postmodernist prophecies of ‘the destruction of meaning’, users of music video and other popular cultural forms continue to derive pleasure from the production of (albeit temporary and volatile) meaning induced by these forms (cf. Straw 1988, Frith 1992).

In spite of the emphasis placed here on textual-structural analysis, it would seem unwise to deny that there are limitations to the explanatory value of such an analysis of music video. It appears equally unwise, however, to discard the possibilities of textual analysis before the nature of these limitations has been investigated by way of such an analysis taking all relevant dimensions into account. Syntactical characteristics are ‘objective possibilities’ (Willis 1978, pp. 198ff.), subject to influence by varying modes of perception, contexts of use and other factors; these factors, as well as the ways in which the visualisation of musical structures is shaped by genre-specific cultural codes of expression, remain to be examined in more detail by future research.

Appendix

Figure 1. Bruce Springsteen: Human Touch
Music and images in music video

Figure 2. Bryan Adams: Thought I’d Died and Gone to Heaven

Figure 3. Michael Jackson: Remember the Time
Endnotes

1 An earlier version of this paper has appeared in the Working Paper Series of the Department of Music and Theatre, University of Oslo. Parts of the theoretical discussion are derived from Björnberg (1992b).


3 The relationship between montage and musical beat is also touched upon by Goodwin, who refers this to the conditions of music video production: 'Directors working under enormous pressures of time will ( . . . ) often shoot some material that can be edited at random to the beat, without any great regard for realism or narrative closure' (1987, p. 26).

4 At the level of visual content, Nielsen also connects this with Straw's argument that the fragmentary eclecticism of music video, rather than representing empty pastiche, constitutes a reconstruction of pop music history within traditional pop song forms: 'The relationship of song to visuals is ( . . . ) rather one between the basic demands of form ( . . . ) and the heterogeneity of codes and visual materials held in play by that form' (1988, p. 258; see also Berkak and Ruud 1989).

5 Most of the works cited here have been written by scholars in sociology, film studies, literary studies or mass communication research. Up to the present, the contributions of musicologists to the study of music video have been few and have also shown a general bias towards the perspectives of these other disciplines, this situation indicating the current embryonic stage of development of a structural semiotics of popular music.

6 Straw's (1988) emphasis on the recourse in the music video era to traditional pop structures (see Note 4) seems to attribute a historically relatively stable and constant, and thus relatively vague and unspecified, nature to this potential ('the basic demands of form'); however, in view of the significant range of variation as regards formal construction in popular music it appears more relevant to speak of 'forms' rather than of one general 'form', and to investigate in more detail the potential experiential consequences afforded by this variation.

7 My earlier discussion of this matter (see Björnberg 1992a, pp. 380ff.), while remaining, to my view, theoretically valid, is somewhat too generalised for the present analytical purposes.

8 The presence of such parts would seem to bring popular music in general closer to the epic category (cf. Table 1 and the analysis section); it may also be pointed out that the
regular metre characteristic of most Western music constitutes a constant element of musematic repetition, although on a rather abstract level.

9 The number of repetitions is crucial for this effect. While a large number of repetitions of directional musematic structures reduces directionality, a 'small' number (2–4) often serves to increase it by means of tension/release effects; cf. Tagg’s (1979, pp. 132ff.) discussion of ‘ready, steady, go’-patterns.

10 Jones (1988) distinguishes between ‘mimetic’, ‘analogue’ and ‘digital’ (non-linear) narrative, where television and music video have been moving towards the third; however, these concepts do not appear immediately transferable to discussions of musical syntax (but may perhaps be seen as parallel to Middleton’s concepts, and mine, as regards the ratio of ‘narrative information’ to total information).

11 This may be illustrated by the fact that with regard to motorial flow, the ‘intro’ (i.e. introductory) section of the song actually extends over the entire first 52 measures (cf. Brolinson and Larsen 1984, p. 345).

12 The essence of the song and the video, like of so much of Springsteen’s production, appears to amount to what Berland (1986, p. 44) ironically describes as ‘the oldest tune in the book: big world, little me’. Existential facts do have a tendency to sound trivial; the interesting point, however, is not so much the triviality of the message as the signification potential (and pleasure) invoked by the specific way in which it is structured.

13 Such deviations appear, in the popular music of the last three decades, to have gradually replaced tonal processes as a means of increasing directionality (in a general sense); whereas in a Tin Pan Alley ballad a sectional transition may be emphasised by a series of II-V progressions with chord substitutions, in a rock ballad one, or a half, extra bar is instead added to a symmetrical period (cf. Björnberg 1987, p. 77).

14 This differentiation of realism levels underlines the structural, functional and experiential differences between traditional film underscore music and the music of music video; however, a detailed discussion of these issues cannot be pursued in the present context (on underscore music, see Gorbman 1987).

15 The video is based on the ‘7” edit’ version of the song, with a duration of 3’41”; the CD single also contains a 5’12” version subtitled ‘12’ mix’, and a 6’49” ‘Purple hazed mix’, the latter with a considerably different and less mainstream-oriented harmonic, rhythmic and formal structure.

References

Allan, B. 1990. ‘Musical cinema, music video, music television’, Film Quarterly, 43/3, pp. 2–14
1991. ‘Sign of the times? Om musikvideo och populärmusikens semiotik’ (‘On music video and the semiotics of popular music’), Svensk tidsskrift för musikforskning, 72, pp. 63–84
1992b. Popular Music Form as Narrative Process, Report series Ljudbilder, 1, Department of Musicology, University of Göteborg
Chatman, S. 1978. Story and Discourse. Narrative Structure in Fiction and Film (Ithaca & London)
Forsman, M. 1986. ‘Det eviga nuet’ (‘The eternal now’), Filmhäftet, 54, pp. 4–14
1992. ‘Det gode, det dårlige og det liggeyldige’ (‘The good, the bad and the indifferent’), Mediekultur, 18, pp. 45–56
Genette, G. 1980. Narrative Discourse (Ithaca)
Goodwin, A. 1987. ‘From anarchy to chromakey: music, video, media’, OneTwoThreeFour, 5, pp. 16–32
Gorbman, C. 1987. Unheard Melodies. Narrative Film Music (Bloomington & Indianapolis)
Grossberg, L. 1988. ‘ “You (still) have to fight for your right to party” : music television as billboards of post-modern difference’, Popular Music, 7/3, pp. 315–32
Ihlemann, L. 1992. ‘Fra rockkoncert til musikvideo’ (‘From rock concert to music video’), Mediekultur, 18, pp. 34–44
Kinder, M. 1984. ‘Music video and the spectator: television, ideology and dream’, Film Quarterly, 38/1, pp. 2–15
Larsen, P. 1987. ‘Bortom berättelsen’ (Beyond narrative), Filmhäftet, 55–56, pp. 82–96
Maróthy, J. 1974. Music and the Bourgeois, Music and the Proletarian (Budapest)
Movin, L. and Øberg, M. 1990. Rockreklamer (Rock commercials) (Viborg)
Willis, P. 1978. Profane culture (London)

Discography
Snap, Rhythm is a Dancer. Logic Records, 665 309. 1992