

Silent Discord: Analyzing Film as Music through *Koyaanisqatsi*

BY MATT BACZEWSKI



Figure 1. Bridge Demolition , *Koyaanisqatsi*, dir. Godfrey Reggio, score Phillip Glass, cinematography Ron Fricke, Institute for Regional Education, 1982, 35:32.

When looking at the above image, one sees a collection of objects coming together to create a picture of a bridge being demolished. What if there were a way to see this as an orchestra, a collection of musical instruments taking part in a visual symphony?

To look at an object of study through a lens of analysis not meant for it is provocative and revealing. To take critical tools and a vocabulary specific to one mode of art and attempt to apply them to another is to expose artifacts and to test preconceptions and former notions

associated with their reception. Looking at film through the lens of music theory- viewing film as music- can in this way suggest a novel investigative technique which revisits characteristics of the medium that have perhaps been taken for granted. Musical theory and terminology can be applied to film and, through this interdisciplinary approach, open new modes of conversation that underscore possibilities for other disciplinary crossings.

Elements of Music

The idea of tension and resolution is perhaps one of the most basic and defining principles in music. It can be argued that no music exists without the interplay of these two elements, regardless of how they are found in a work. A piece of music creates tension through any number of techniques before, through an equally vast possibility of methods, resolving that tension. Since the dawn of the twentieth century, new ideas have exploded regarding how to organize this interplay between tension and release, often breaking into experimental territory; until this point, however, dominant conventions had governed Western art music for hundreds of years¹. These are the ideas of consonance and dissonance, and the use of harmony and melody- moving through consonance and dissonance- to achieve conflict and resolution.

¹ Describing the centuries of pre-20th century Western art music as “conventional” and lumping them all together is not meant here to oversimplify or demean the music as fitting neatly within one classification and understanding. Throughout these years, there were countless advances and iterations of theory and musical styles- some considered revolutionary at the time- that should not be ignored. It is only in light of the experimentalism of the twentieth century and the radical changes that came to redefine music in many ears, oftentimes disregarding past conventions of harmony, melody and rhythm altogether, that a description of pre-twentieth century theory as conventional can be appropriated.

To look at harmony and melody, one must first understand pitch. Pitch is essentially the frequency of a sound. It can be high or low or anywhere in between. Relevant to this is timbre, which is the sound “quality” of the instrument being played. Pitch is represented, in standard notation, by notes placed in different positions on the musical staff; the distance between one pitch and another is known as an interval. These individual pitches can be played alone or can be combined into a chord, a collection of two or more pitches. A chord can be classified a number of ways according to the intervallic relationships of the contained pitches (examples include major, minor, and diminished).

Chords and pitches are the building blocks of harmony and melody. Melody is, conventionally, the featured line, or what the composer perceives as the focus of the piece. Harmony is the music, most often played as chords, that “supports” the melody; it is the background that provides context to foreground melody and allow it to flourish. Harmonic structures will vary greatly depending on genre, time period, and a composer’s own creative intentions. These structures arguably are always based on the feelings or experience of tension and resolution.

Often, according to Western art tradition, music will start with a chord or sonority that sounds relatively stable. This feeling of stability, of solid ground, can be called consonance. As the piece proceeds, a composer uses various compositional techniques within a harmonic progression to create a feeling of tension and movement away from that original “solid ground” or stability. This tension, an effect produced by sonorities which are unstable, can be called dissonance. Dissonance often creates intrigue in listeners and is responsible for effects similar to

conflict within a narrative. Tension is frequently “resolved” by returning to consonance, which satisfies the listener as the problem, introduced by dissonance, is settled.

Another important element of music is rhythm. Because music exists in time, it is necessary to be able to measure and denote this temporality. Once again, the twentieth century has produced a number of avant-garde theories about rhythm and its use; until then, much like ideas about harmony, rhythm was dealt with somewhat conventionally.

Rhythm is a demarcation of how music exists in time, or how it occupies space in the medium of time. A composition is divided into measures, which are then divided into beats. The beat is the internal pulse of the music, which can speed up or slow down either gradually or immediately; this speed of the beat is the tempo. A note can then be sounded for any duration that divides or multiplies the beat. Rhythms in a piece can be simple or can develop into complex patterns of notes. Because music exists in time, however, rhythm is always present.

Film as Music

To look at a film through the lens of music theory could be done in two different ways. One way would be to view a single frame as an entire musical composition, identifying the individual components of harmony, melody and rhythm as they exist in the image and create a whole piece. Another approach would be to view the entire film as a musical composition and identify each single frame as a chord or sonority that exists as a harmonic unit within the composition. Theoretically, this possible translation of film and music could be tested on any

moving picture. Reggio's *Koyaanisqatsi*, though, provides an especially interesting test case because of the film's emphasis on composer Philip Glass' score. Throughout the production and editing processes of the film, Reggio and Glass frequently worked in conjunction in order to more effectively synchronize the two. This gives the film an inherent musical quality and flow and makes it an effective starting point for such an investigation.

All of the elements of a musical piece should be able to be located within a single frame of a moving picture if it is to be understood as music. A frame from *Koyaanisqatsi*, in which a bridge is being demolished (Figure 1), offers an example in which to demonstrate such a determination and dissection.

Timbre can be taken into consideration in film by looking at the qualities of the film itself. The definition of lines, the brightness and contrast, the use of color, and the "grittiness" of the film could all factor into timbre. One could look at the bridge having a cold, sharp timbre like that of a piano, or at the explosions as having a reedy, gritty timbre like that of a saxophone; even the sky could be thought to possess a smooth, light timbre like that of a flute. This could be applied to objects in order to reveal a symphonic experience of moving images.

An initial question to address is if and how a still frame can be said to reflect tension and resolution. Objects, acting as chords, could possibly exist within and be identified as having dissonant or consonant relations. One way to apply these terms would be to view each object in the image as a chord and then look at the relationship between the objects to understand their function as a chord. Looking at the bridge, three sections can be seen- the intact left side of the bridge, the falling center, and the intact right side. The left side is still upright, still stable. Its supports have not fallen, there have been no explosions in the region, and it is supporting itself.

The same can be seen on the right side, a particularly stable section of bridge. In between the supports, however, the bridge has been demolished and is very much unstable, falling through the air.

In this rather architectural talk of stability and instability is a key to the tension and resolution of music. Both sides of the bridge might be understood as points of consonance, solid ground where rest can be found. They are not in flux. They are, maybe, somewhere that the eye can stop and feel relaxed, much as the ear could spend time mulling about a consonant sonority because there is no perceived need for resolution; it is already stable. Compared to these two sides, the center is a point of tension. Nothing is settled as the section of bridge is coming apart and falling through the air. This unsettled space might be articulated as dissonance—a sense of instability that creates tension in the image. This bridge then, in a sense, exemplifies the elements of harmony that create music, a place of stable consonance interrupted by the tension of dissonance which is then resolved by a return to consonance.

A problem that could arise through such a method of translation is how exactly to determine stability in the image. In one way, an analysis of the image could rely on knowledge of the image's subject matter. One person may understand that an explosion is taking place and disrupting the center section of the bridge, understand that the center section is not settled but rather in flux, while the rest of the bridge seems unaffected by the explosion. Another person might experience instability on the left and right sides because bridges are, in themselves, potentially precarious. In other words, each viewer brings unique experiential circumstances to the perception of an image which may result in wildly different, even contradictory interpretations of what counts as stable, consonant, or discordant. While this may seem

problematic to the possibility of using circumstantial knowledge of the image's objects as a way to determine their functional roles, it is surprisingly similar to how a listener perceives music. Different ears may have different reactions to dissonant chords, depending on their own experiential circumstances regarding sound. An ear accustomed to hearing dissonance may have a different response to the question of aural instability than another. In this respect, then, the determination of stability as a subjective one in film is comparative to the individual's determination of stability in music.

It is possible, however, to also attempt an approach that could reveal itself more objectively, much the way that by looking at a chord written on a staff a musician could perform a more empirical measurement of its quality. A possibility would be to look at the structural characteristics of the visual object. The left side of the bridge as a visual shape can be seen to be constructed from defined and simple lines, repeated rhythmically in an even pattern; the same can be said for the right side of the bridge. This attention to geometric attributes lends itself to a more universal interpretation of its structural integrity. In contrast, the falling center section is comprised of interruptions in the form of multicolored, amorphous clouds of smoke and is bent near the right side. Compared to the rigid patterns of the end sections, this could almost be seen as a visual anarchy. The structural integrity of a visual object that is anarchic, whose structure is continually and somewhat rhythmically interrupted, and its disruptions could be said to correspond with dissonance, an explosion of noise. The geometric regularity of the uninterrupted side may correspond with stability and consonance. The straightness, definition, and pattern of lines and shapes can be measured and used to understand this harmony and dissonance.

It is valuable to explore the musicality of film through translations of the musical concept of harmony because the chordal quality assigned to an object could reveal a large amount of information. Looking through the possibilities and methods previously discussed, one might be able to understand that an object is experienced as dangerous or unstable, is visually abstracted, chaotic or interrupted, and provides tension to the image, merely by knowing that it is considered “dissonant.” This could be an exciting way to convey a large amount of information in much less language.

Identifying the melody in a single image seems to be a harder task. If this approach is dependent on viewing the objects within a frame as chords, then there is little room left to identify a melody. Perhaps the object that is most focused on could constitute the melody, though it would already be used as a chordal object. Maybe an approach which favored the object with the greatest illusion of motion or action could be more successful, though a determination of this would be hard to deliver.

Rhythm is an essential aspect of music and potentially troublesome when translated to the still frame, considering that it does not exist in time as music does. Perhaps, rhythm could be the visual arrangement and size of the objects in the frame. This approach is conceived by looking at the relationship between space in a frame and duration in a musical piece. The still frame exists in a different medium, on a physical plane, while music exists in time. If, however, one were to view rhythm as the amount of time a musical event occupies, similarities appear. Rhythm is a measure of the amount of the art’s respective medium that an object occupies; in music this means the amount of time that a note or chord occupies, time being the medium,

while in a still frame this means the amount of space that an object occupies, with physical space being the medium.²

In Figure 1, the bridge poses an intriguing “rhythmic” picture. The edges could be seen as an undisturbed sound, having a relatively large space uninterrupted by any other objects. The center section’s rhythm, though, could be seen either as being interrupted by the clouds of smoke disrupting its visible continuity or as possessing a large rhythm “sounding” underneath the staccato rhythms of each cloud of smoke. In this case, the center could be rhythmically dense and complex, due to the interplay of larger and smaller rhythmic values, while the edges remain rhythmically simple places. Perhaps, the background of the skyline, too, could be said to be rhythmically complex because of the diversity of the buildings and their interruptions into the sky above. When these rhythms are combined, there would be room for a musical type of interaction between varied rhythms due to size and interruption.

Identifying a single frame as a piece of music reveals certain characteristics of film in novel ways. One problem with the approach is that it is difficult to pinpoint a progression in the viewer’s perception; that is, it is difficult to know in what order a person viewing a film perceives objects within an image. It could be that they see the image from left to right, as if reading it, or vice versa; the image may be perceived in a vertical progression as well. Possibly the image is even perceived all at the same time- with all these different possibilities of perception it would be hard to delineate how musical resolution comes into play as a

² Medium could be viewed in five different ways- as a vessel, as a support, as a cultural domain, as criteria for differentiation according to materiality and transmission channel, or as semiotics. See Marie-Laure Ryan, “Narration in Various Media,” in *Handbook of Narratology*, ed. Peter Hühn, *Narratologia : Contributions to Narrative Theory* 19 (New York: Walter De Gruyter, 2009), 263–81.

progression from an area of tension because it would be hard to delineate the forward movement of the perception of the image itself, if in fact there even is forward movement.

It may be vital to also explore an approach which utilizes film's existence in time, an important quality which it shares with music, with or without cuts. An analysis done in this way might focus on the frame as a chord with its own qualities of dissonance and consonance, instead of as a full composition. This frame leads to another chord, continuing through sequences of tension and resolution throughout the film. Each frame becomes a piece in the harmonic progression and structure of the work.

The sequence in *Koyaanisqatsi* from which Figure 1 is taken can be used to explore this idea of chordal frames. The sequence (Figure 2) shows a number of buildings and other structures being demolished. Each time the structure is shown mostly intact, followed by its explosion and demolition. When one explosion, one area of tension, occurs and is let play, it is then resolved by the appearance of a new building, which in turn falls and becomes an area of tension in the piece. The frames selected are an excerpt; each frame depicting an explosion on the left is followed immediately in the film by the frame on its right.





Figure 2. Frames (as chords) supplying tension. A frame of tension followed immediately by the next frame of resolution. Demolition sequence, *Koyaanisqatsi*, 35:20- 36:26.

A determination must be made of how a frame becomes qualified as either consonance or dissonance. In music, this quality depends on the intervallic relationships between the pitches contained in the chord. Translated somewhat literally, this could mean for film that the quality of a frame as a sonority is determined by the condition of the relationship between the visual objects within its boundaries. Here, as before, these relationships should be looked at in a way that allows for the subjectivity of the viewer and a more objective, empirical approach.

Perhaps the first method for determining these relationships between objects in a frame, acting as pitches within a chord, is to look at how objects interact with other objects. If objects are in conflict within context or with other objects within the frame, it could be denoted as a chord supplying tension; if objects are interacting positively within context or with other objects within the frame, the image could be denoted as a chord supplying resolution. Again the viewer

must put into effect his or her own experiential histories and circumstances in determining positive and negative interactions, just the same as discord and consonance in music may depend on the listener's prior experience of such phenomena.

Looking at the first frame in the sequence, one notices the overall sense of disarray in the scene, caused by the explosion and demolition of the building. Part of the building is visible as it crashes down to earth through billowing clouds of dust invading the landscape, moving in indefinite shapes and sizes; entropy abounds, and disorder breeds an atmosphere of tension. The clouds look alien and disruptive to the landscape, and the falling building does not seem to be designed for such an interaction with the ground. This frame, then, could be seen as one which bears disorder and tension, which would imply a discord, a dissonance between its objects or pitches. The next frame, to its right, is an intact, freestanding building. In this frame, there is order and positive interaction between the objects that make up the image. The buildings appear to fit neatly in place where they belong, in well-defined shapes and angles. This is an undisturbed scene; no object is in conflict with another object. The frame thus may be determined as stable, a place of resolution after the tension of the previous frame. The same process could be tested on each of the subsequent pairs of images from the sequence.

If viewing film through the lens of music theory is to be successful, it is important to have a more objective approach in addition to the previous one. Perhaps the quality of the frame could be determined through an analysis of the visual structure of the image. A frame which possesses a well-defined, rigid, orderly structure may be a frame which provides stability and resolution, while one which possesses an abstract, unruly structure may provide a feeling of tension. Considering the first frame of the second pair, the viewer might conceive of an

apparent lack of order in its visual structure. Shapes are colliding into and on top of each other, patterns are difficult to discern, and the crashing towers have lost their previously rigid structures. Visually, the relationships between shapes are disorderly. Moving on to the next frame in the film, returning to the bridge from Figure 1 before it has exploded, the viewer finds a stable bridge, structured in a repetitive pattern with carefully delineated lines and unyielding steadiness; the skyline in the background is nearly parallel to the bridge. Visually, the relationships between shapes are orderly and geometrically sound. This could be the return to a consonant chord, a solid ground, after the tension of the previous frame.

When looking at the whole film, melody may be easier to translate from music to film. An appropriate approach may be to use Roland Barthes' idea of *studium* to determine what is meant to be focused on in the frame. Barthes, in his book *Camera Lucida*, theorized that an image contains a *studium*, which he defined as an "application to a thing, taste for someone, a kind of general, enthusiastic commitment, of course, but without special acuity."³ He goes on to say, "To recognize the *studium* is inevitably to encounter the photographer's intentions..."⁴ In this sense, the focal point is where we encounter the possible intentions of the filmmaker, or what he or she is trying to draw out and present to the viewer. In music, this focal point would be the melody, the melody which appears to be intentionally brought forward and focused on by the composer or musician. In the first frame of Figure 2, this could be the crashing building, just above the ground; in the second frame, this could be the building which is neatly centered in the frame. Melody could possibly become a progression of these *studiums* as they change from frame to frame.

³ Roland Barthes, *Camera Lucida*, trans. Richard Howard (New York: Hill and Wang, 1981), 26.

⁴ *Ibid.*, 27.

Rhythmically, looking at the entire film could mean focusing on the editing of the film, mainly on the film's cuts. This seems to be an almost literal translation of musical ideas, as well as a descendant of the ideas of Dziga Vertov and Sergei Eisenstein. These two directors and theorists were part of a generation of Soviet filmmakers who created and fostered the idea of montage editing. This form of editing stressed both the use of association between images, instead of direct action, and on the musical rhythm of editing, or the tempo of cuts. In his writing "We: Variant of a Manifesto," Vertov discusses the musicality of this style, writing:

The meter, tempo, and type of movement, as well as its precise location with respect to the axes of a shot's coordinates and perhaps to the axes of universal coordinates (the three dimensions + the fourth- time), should be studied and taken into account by each creator in the field of cinema.⁵

So Vertov felt that rhythm was an essential aspect of filmmaking, and that the tempo of cuts was necessary to establishing a certain cinematic experience.

Koyaanisqaysi is an interesting example to look at in this respect because it was edited in conjunction with Philip Glass's original score. In his article "Sight, Sound, and Myth Making," Mitchell Morris explains:

Glass would often start with very rough cuts of footage- sometimes only single images- and write a movement of continuous music. Reggio would then often cut the film to match the music, after which the two would adjust the timing and coordination to their mutual satisfaction. The final shape of the film is thus a series of musico-visual movements variously linked together in ways that may seem ritualistic.⁶

⁵ Dziga Vertov, "We: Variant of a Manifesto," in *Kino-Eye: The Writings of Dziga Vertov*, trans. Kevin O'Brien (Berkeley: University of California Press, 1984), 8.

⁶Mitchell Morris, "Sight, Sound, and the Temporality of Myth Making in *Koyaanisqatsi*," in *Beyond the Soundtrack: Representing Music in Cinema* (Berkeley: University of California Press, 1984), 122.

The film had already been edited in a way that dealt with specific rhythm. Perhaps this method of editing as musical rhythm is already well evidenced and important to validating a musical approach to film, as rhythmic editing is a previously accepted form of filmmaking.

What is interesting about the editing of the sequence shown in Figure 2 is that each specific explosion is never allowed to resolve itself before a cut is made to another, intact building that then explodes. This rhythm creates tension in itself, while not interfering with the tension and resolution necessary for musicality. This is a play on the aspects of tension and resolution itself, where the crashing building is not resolved by returning to its previously stable state but rather by introducing a new object of stability, a new consonance different from the previous one.

Another possible way to look at the rhythm of a film could be to analyze the tempo of onscreen action. This could be especially pertinent to *Koyaanisqatsi*, in which many shots are either sped up or slowed down in order to create unique visual effects or emphasis. The demolition sequence is one such spot. The buildings subject to explosions do not fall at the same rate, nor do they fall at a realistic rate. The tempo is slowed down in order to emphasize the falling of the buildings, and to give the spectator a longer period of time to experience this falling. Another place where a realistic representation of time is eschewed in favor of visual effect is the highway traffic seen in Figure 3.



Figure 3. Traffic at night, *Koyaanisqatsi*, 1:02:36.

In this shot, footage of highway traffic at night is dizzyingly sped up in order to create a hypnotic visual effect. This is done not through the editing and cutting of footage, as it is in other parts of the film, but through the manipulation of the footage's tempo.

Possibly the idea of rhythmic editing and the idea of rhythm as the speed of action in the film can be applied at the same time, with editing taking the role of rhythmic duration while speed of action is translated into the tempo. A musical theory of rhythm thus applied to film does not seem to contradict the medium; in fact, because of the temporal nature of each medium, this theory seems especially fitting.

The application of this analytical framework is not limited to any one film or piece, but rather can be utilized on a more overarching scale. *Koyaanisqatsi*, being a noteworthy showcase for the interactivity of film and music, provides an appropriate starting point for an exploration of this framework; by no means, however, would this combination be necessary in order to attempt such an analysis. This method of viewing film as music is not reliant on a theatrical score or sound in general. The analysis is concerned with aspects of the film medium that are

not reliant on the presence of the aural, and it is therefore possibly applicable to any piece working within such a medium.

Analyzing film through music theory may provide critical new insights into the medium, or perhaps a new way of discussing and viewing film. There are a number of possible advantages to this. One may be a way to lend support to new investigations into the shared properties of different media, an exploration that could be continued with other mediums. A musical analysis could possibly provide an informative new way to talk about film, as discussed before, including new terminology and concepts borrowed from music theory to either enhance current discussion or dramatically alter it. Perhaps these ideas could even be useful to filmmakers, if they find them to imply new ways to make and edit film. Viewing film as music could, if taken seriously, provide vital perceptions of film and its ontology, as all interdisciplinary methods of study inevitably offer a new magnification through a new lens.