*Performing Performances: Some considerations on the role of the filming and recording crews in classical music videos.*

Gaia Varon – University of Bologna – gaiavaron@gmail.com

Nowadays, it is common practice to 'watch' symphonic music on a screen. DVDs, Music TV Channels, old videos found online, concerts in live streaming, have become part of the daily life of music lovers. New videos are produced every year, and watching older ones becomes easier and easier. What is the nature of a classical music video? What kind of relationship does it establish with the score, and with the performance it displays? How faithful a document is it of the performers' choices? And what is the impact of the choices of the filming and recording crews? This paper presents the provisional results of an ongoing research, aimed at outlining and analyzing the structural mechanics of the screen version of a performance, particularly the relationship between the screen version and the score, and that between the screen version and the performance.

The research has a purely pragmatic starting point: for most music lovers today, and especially for symphonic music, the object through which music is experienced, is certainly not the score, i.e. certainly not an ideal art work as it is fixed and guaranteed by the score. In most cases, it is neither the actualization of that score in a specific performance experienced in the concert hall. The most common aesthetic object is most often a specific performance of the past as it has been fixed once and forever through recording, whether audio or video. More and more often, the live experience of a concert is

also a mediated one – whether through radio, television or the internet. The research focuses on the video portion of these daily aesthetic objects. It also takes a few assumptions as starting points. First, filming music implies a specific authorship, comprising all the individual contributions of the team components. Second, the screen version of a performance can be considered a new, specific object, in which three different layers of authorship coexist: composition, performance, screen production. I call the level of screen production the 'third authorship', and I assume that analyzing the third authorship layer is possible, legitimate and interesting. My interpretations are grounded on personal professional experience in music film making, and on empirical research, consisting in direct observation of, and interviews with the crews during filming, broadcasting, and recording operas and concerts at Milan Teatro alla Scala.

The 'third author' of music filming most often includes an audio recording team and a filming crew, and the two groups may work quite autonomously from one another. The recording crew necessarily includes components with high skills in both music and technical matters; there is maybe one person combining both, the sound engineer, or, as it is more common in Italy, two different people, the music expert and the chief technician, working closely together. For the filming, most directors also rely upon a music consultant. The audio and video music consultants – rarely, there may be one person only, contributing to both the audio and the video recording – preliminarily work alone on the score and make a plan of the oncoming recording. From the perspective of my research, this is the most relevant phase of the work: this is when aesthetic, cultural, even ideological choices are made, which will then be combined with material decisions, such as how many microphones to set up, what type of microphones, where they should be positioned, what should the camera focus on, when to show the conductor and when the soloist, etcetera. In other words, the third author identifies specific bits of meaning in the music, makes a selection among them while establishing a hierarchy of their relevance for what in this perspective we may call the alleged audiovisual translation of the music piece and/or of that specific performance of the piece. In my research though, I do not focus on the material process of this phase, but I analyze the final product, the video recording, aiming at bringing to the fore the results of the decisions that were presumably taken during that phase. I will today offer a few brief glimpses in the kind of analysis upon which my research is based.

There are obviously several issues at stake in the decisional process that leads to an audiovisual recording; and I will today focus only on two of them: one, the impact of the video and audio recording choices on the alleged spectator's perception of one specific music element; two, how the shooting and the editing choices can clarify, enhance or conceal the formal organization of a piece.

The first issue concerns how the image can be used to enhance the presence of a specific sound element and thus make it more easily recognizable; this is mostly done through close-ups on the source of that specific sound element, but the results of such visual strategy also depend on the sound recording choices. Let us briefly go through some examples.

In Beethoven's *Piano concerto No. 3*, the *Rondò* begins with a ten bars theme played by the piano; the same theme is then echoed by the oboes mostly doubled by the bassoons and with a sustaining line played by the horns. While the woodwind play the theme, the piano plays a decorative line, a light passagework for both hands. In the first ten bars the theme is thus clearly audible, while in the following ones, when played by the woodwind, it competes with the piano part.

In different video recordings, the 'presence' – in terms of sound – of the two elements may be slightly but relevantly different: the difference may depend on the performers' choices, but also on the choices of the recording crews: the soloist has his or her own microphone, which may be more or less close to the instrument, giving it accordingly a more or less present sound; and the same is true for the microphones set up for the orchestra and for the individual instruments within it. During a recording, and again during the editing, several technical devices maybe used to increase or decrease the volume of a single microphone.

I have analyzed several recordings, and I will now briefly discuss three of them: Gilels, Berglund, 1984; Zimerman, Bernstein, 1989; Lewis, Bělohlávek 2010. The sound recordings are recognizably different: in the Gilels version, the piano is always very prominent (probably the microphone was close to the instrument) and it prevails over the woodwind during their playing of the main theme; Zimerman is well balanced; in Lewis, the theme is always in the foreground.

The videos on the other hand are very similar.

**Gilels, Berglund, 1984**



**Zimerman, Bernstein, 1989**



**Lewis, Bělohlávek 2010**



In the three versions, as in many others, the same tool – a shot of the woodwind – is used to enhance the 'presence' of the woodwind, to make them visible and thus to make them more 'audible'. Such homogeneity of filming strategies is in this example quite reasonable: in a passage like this one, the music syntax and the film syntax perfectly overlap in creating and giving evidence to the 'question and answer' character of the passage. But the

three sound recordings, as we said, are very different, and the audiovisual combinations function thus quite differently: in Gilels, where the piano sound is always predominant – not only in volume but also in terms of sound space – the image of the woodwind functions as an aid, a didactic help, we could say: it directs the spectator's attention to an element which the 'video third author' – the director, his music consultant – considers important. In terms of sound, the element slips in the background, the image brings it to the fore. In Zimerman, the video enhances one after the other the two elements that are well balanced in the audio; it brings the spectator's attention to the music syntax and to the theme passing from the soloist to the orchestra. The Lewis BBC recording, on the other hand, offers a perfect audiovisual correspondence: the video literally illustrates what we hear.

Let us now focus on another instance, where the recording choices fulfill a slightly different task: within a mix of several sounds, the range of the third author's possible choices goes from identifying, selecting and bringing to the foreground one specific individual element of the mix, to aiming at a fully blending mix. I will briefly discuss this issue focusing on a passage of Beethoven's *Symphony No. 7*.

A distinctive feature of the *Poco sostenuto* at the opening of the Symphony's first movement is a marked alternation of two elements: some chords played *forte* by the full orchestra, alternating with brief, slow melodic fragments, *piano*, played by the woodwind (oboe, then clarinet, then flute). Further on, at bar 10, there is again an alternation, now between ascending scales played by the violins, *staccato*, with some sustain of the other string instruments, and brief, sweet (*dolce*) chords at clarinets and bassoons; the brief sequence begins *pianissimo* and increases up to a *crescendo*, eventually ending *fortissimo* at bar 15. The scales are repeated three times, and at the third time, at bar 14, some of the woodwind double the violins (the flute doubles the whole scale, clarinets and oboes only some fragments after which they keep playing repeated notes, with bassoons and horns joining in). The audio recording may aim at keeping the different elements – the sound of each individual instrument – clearly distinct and easily distinguishable one from the other; or it may try to achieve a full blend. And in different recordings we are actually presented with different choices.

Let us focus on bar 14, where the flute comes in doubling the violins: in some recordings the flute is clearly in the foreground, while in others it blends with the other instruments. As for the video, quite obviously, there may or may not be a shot of the flute. So, here again, we will find various possible audiovisual combinations. Out of the several video recordings I analyzed, I will briefly discuss six examples.

**Kleiber, Concertgebouw Orchestra Amsterdam, 1983.** In the audio recording, the timbres are kept well distinct from one another and the listener can easily detect the sound of the flute when it comes in. On screen, after some small exchanges of brief shots, the image of the conductor is kept during the rest of the sequence until bar 15; we do hear, but we do not see the flute when it comes in.[1]

**Rattle, Europa Konzert, 2008.** The sound recording seems to aim at a perfect blend; the flute does not stand out, in other words we do not detect its specific sound within the sum of all the instruments playing. On screen though, after a fairly long zooming shot on Rattle, we have a series of close-ups on individual instruments: violins, oboe and clarinet, violin again and eventually the flute.





---

[1] http://www.youtube.com/watch?v=s1qAWcd4rr0

This way of filming is a good example of what I call in my research the 'standardized approach', i.e. an approach that aims at a good visual balance between all the ingredients – shots of the conductor, shots of the orchestra, details of individual instruments. The details have a very clear, somehow 'didactic' function: they bring the spectator's attention to a specific sound element by showing its source. Such 'informative' details are a basic feature of the standardized approach. This approach is today the most widespread and is used in many recent films of Beethoven's *Symphony No. 7*.

**Harding, Filarmonica della Scala, 2008.** With slight differences in framing, the filming strategy – the objects and types of the shots, the cuts – is the same as in the Rattle's version examined above.



And here again, the visual part has no correspondence in the sound recording: we see the flute but its sound does not stand out. The image of the flute functions in both these films – as it did in the Gilels example – as a didactic tool, bringing the spectator's attention to a sound element which would otherwise go unnoticed.

**Abbado, Berliner Philharmoniker, Santa Cecilia, Roma, 2001.** We are confronted again with the standardized approach – although with fewer shots; the sound recording though is analytical and the timbres are kept well distinct. The result is a perfect audiovisual correspondence: we see the flute while we distinctly hear its sound.



**Karajan, Berliner Philharmoniker, 1971.** The image of the conductor appears on screen during the whole sequence; the flute sound is blended with that of the other instruments.



Summing up, in the five films examined we met with all the possible combinations: we see the flute on screen while its individual sound is clearly detectable within the sum of all the instruments playing (Abbado); we see the flute on screen, but we cannot detect its individual sound within the blend (Rattle, Harding); we do not see the flute on screen, but we detect its individual sound within the sum of all the instruments playing (Kleiber); we do not see the flute on screen, and we cannot detect its individual sound within the sum of all the instruments playing (Karajan).

Each combination affects our perception of the music. Seeing the flute and hearing distinctly its individual sound is the most informative strategy, it guides effectively the spectator's attention towards one specific musical element, implicitly bringing in an analytical – albeit most often unconsciously so – way of listening. The consistency between the video and the audio recording strategies allows the standardized approach to function properly and effective. Seeing the flute without distinctly detecting its sound is probably also meant to be informative – a didactic tool, as we said, telling us that a specific element is worth attention; but in fact it risks functioning almost as a mere decoration: we see the flute, but the sound does not support the perception of a particularly relevant role of the flute in that specific bar. The lack of correspondence between the video and audio recording strategies weakens the audiovisual combination. In the other two combinations, where we do not see the flute, the screen brings the spectator's attention elsewhere; in most films of Beethoven's *Symphony No. 7*, when at bar 14 of the *Poco sostenuto* we are not presented with a shot of the flute, we see the conductor, or a wide shot of the orchestra with or without the conductor. And the effect of the shot is mostly emotional; it brings the spectator's attention to the sense of the wholeness of the musical performance, conveying either the sense of the live event, or the power of conducting, or the power of the music itself.

There is at least one film where a different and interesting choice is made. There are two filmed versions of Beethoven's complete Symphonies conducted by Herbert von Karajan; the first one dates to the beginning of the 1970s, and different directors are involved – quite often with an 'artistic supervision' of Karajan; for the second one, Karajan himself directed all the films.[2] Although different in other ways, the sound recording of the 1983 film of *Symphony No. 7* resembles that of the already mentioned 1971 film: in both versions the recording seems to aim at a perfect blending; at bar 14 the flute does not stand out and its individual sound is not detectable within the sum of all the instruments playing. The visual solutions though are different: as we know, in the 1971 film the camera stays on the conductor during the whole sequence. In the 1983 film, the image of the

---

[2] This superposition of roles – the conductor acting also as the 'third author' – is an exception. In all the other video recording examined here, as in most video recording, the roles are distinct.

conductor alternates with details of individual instruments and shot of small groups of musicians; at bar 14, a detail appears which is not the flute:

**Karajan, Berliner Philharmoniker, 1983.**



At the moment when the flute comes in, we are presented with a shot of the double bass; the spectator's attention is thus brought not where it most frequently is – the highest line, the most naturally prominent one – but to the lowest register – the bass line. The choice is decidedly unconventional and very interesting: both the flute and the double bass are there, in the score, in the music. But Karajan does not direct the spectator's attention to the figure standing against the background, i.e., to an element easily standing out; he directs the attention to what we usually perceive as the background, the lowest line, which is in a way the generating line of the music.

In terms of filming strategies, the tool used here – a shot of the source of a specific sound element in order to bring to it the spectator's attention – is the same we had already seen in other instances. But the musical thinking that presides over its use is a very different one.

Strategies aiming at orienting the spectator's attention may also be used on the horizontal plane; in a film, necessarily, images line up, bringing about their own sequence, their own 'story' we may say, which is superposed to and intertwines with the music line. Let us now move to the second issue I would like to discuss in this presentation, i.e. how the shooting and the editing choices can clarify, enhance or conceal the formal

organization of a piece. As an example of how to analyze it, I will here focus only on one precise and simple element.

In the first movement of Beethoven's *Fifth Symphony*, a brief call of the horn sets the first and the second theme apart. It is a very effective element of transition, because it functions at the same time as a sort on an echo bringing the first theme to an end, and as the opening of the new theme of which it introduces the key, E Flat major. The precise point of divide between the two themes is not unanimously identified: one of the first commentators of the *Symphony*, E.T.A. Hoffmann describes the call of the horns as a new entry of the first theme and writes of the second theme as being introduced by the violins; other analysts though consider the call of the horns already as a part of the second theme. Implicitly, if not consciously, a director takes a stand (and it may be indifference) when he decides where to cut from one shot to another. Here the issue of sound is extremely simplified, as the horns are the only playing instruments, and the choice concerns the video only.

The most obvious choice is to cut to the horns at the beginning of the call, and then to the violins after it has ended, i.e. in terms of (implicit) analysis, setting the call as an isolated transition element. And it actually happens very often:

**Karajan, BPO, 1957.**             **Klemperer, NPhO, 1970**             **Bernstein, WPO, 1977**



**Karajan, BPO, 1983.**             **Muti, Scala Phil., 1998**             **Abbado, BPO, 2001**

Here again, we have in audiovisual terms a perfect correspondence – redundancy, we could say. But I would like to move towards my conclusions by mentioning two versions where the horns are not granted such an ordinary shot.

The first is part of the NBC telecast of Toscanini conducting the NBC Symphony Orchestra, on March 22, 1952. It was directed by Kirk Browning, then very young, and later one of the most prolific and well–known TV directors. The first shot:
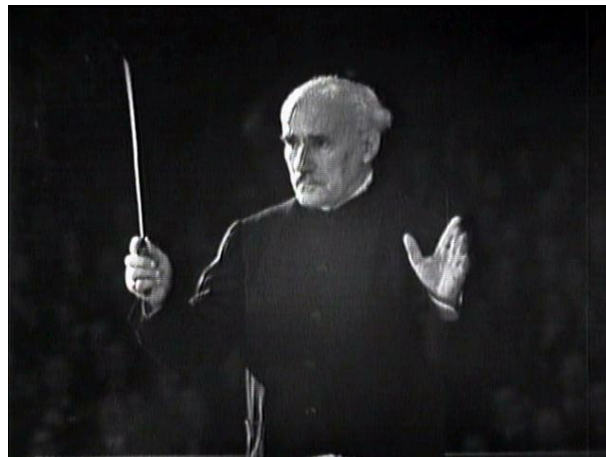


is kept through the call of the horn; a change – the first cut from the beginning of the film – comes in with the theme played by the violins.

In terms of (implicit) analysis, we could say that Toscanini's director, Kirk Browning, has taken Hoffmann's stand, considering the horn call as part of the first theme. In this film, though, not only are the horns missing here, but there is no detail whatsoever; actually, during the whole first movement we never see a detail. Toscanini is there all the time. This is not a necessity, it is a choice. The image of Toscanini with which the film opens (see above) has been compared to a Rembrandt painting, because of its darkness and of the white element standing out of it. But there is also something religious: Toscanini is represented here as a minister of cult. In this image a whole conception of music, and of a contemplative, absorbing way of listening is implied.

Let us now look at the film of Beethoven's *Fifth Symphony* conducted by Karajan in 1966; it is part of a special project for a series of music films with Karajan conducting and the well-known French film director Henri-George Clouzot. Black and white here is partly a choice, a stylistic feature. A strong contrast of black and white was a central element of Clouzot filmic style. The set and the shooting converge in conveying the sense of a great sound machine, with the conductor as its engine. And again, we can focus on one detail to grasp the essence. During the first minute and a half of the film, there are four shots; first, we are presented with Karajan alone giving the starting cue (six seconds), then with the conductor at the center of the whole sound machine (thirty–three seconds); when the second theme comes in, we see again Karajan among the players, now shot from the left side in a somewhat 'warmer' image (fifteen seconds).

No horns, no detail whatsoever. Then something happens which is unexpected: the repetition of the exposition starts with a shot totally unusual for a concert telecast.



During the first minute and a half of the film there are no details; we do not see the horns or the timpani that can be found in most films of the *Fifth Symphony*. These details were missing in the Toscanini's video because they would not serve (and maybe even spoil) Browning's purpose of inducing a contemplative, somewhat religious hearing attitude in the audience. In Clouzot's film the reason for the absence is different: any detail would spoil the effect of the detail of the conductor's hands, the first detail to come in in the film. And this detail tells a different story: it is not Karajan's story anymore; it is now Clouzot's story. These powerful hands against the strongly contrasted background use specific cinematic features to convey a sense of drama which is closer to Clouzot's films than to the alleged Beethoven conception of the 'Symphony of Fate'.

Conclusions.

My research aims to unravel the long and complex work implied in producing the screen version of a concert and to bring to the fore the contributions from people with different and specific abilities and qualifications. Each of them, either separately or during team work, brings about technical, material and cultural decisions. And all these small choices merge into one specific and unique final object. The audiovisual recording of a musical performance is thus a multi-authored work, and the third authorship is never neutral, but necessarily superimposes its own reading of the piece and the performance. When we watch Daniel Harding conducting Beethoven's *Seventh* at La Scala, we must be aware that we are not experiencing that particular performance, but a specific reading of it, i.e. the specific 'overscore' of the third author – the performance of the performance, so to say. A better understanding of the series of small creative acts that line up during the production of each of these material objects in which the performance of a performance of an original artwork is fixed is an important component of any reflection on today's musical landscape. It will help both scholars and musicians better to understand performances, both theirs own and others'. It is also an interesting way to approach the issue of whether this artifact may be granted the status of a new and specific form of the artwork. Or a new artwork altogether perhaps?