

Michel Chion in *Audio-Vision* and a practical approach to a scene from Andrei Tarkovsky's *Nostalghia*

Josep Torelló
Jaume Duran
Translated by Raúl Gisbert Cantó

Audiovisual analysis of the music in relation with the image, within a general framework of cinematographic studies, may not yet have sufficiently explored how music and sound influence the creation of an audiovisual discourse. It is plausible to state that music, when it appears on the screen, when participates in the scene in any of its audiovisual forms, influences, from an ontological point of view, the creation and definition of the filmic diegesis; music has a decisive influence on the creation of the audiovisual discourse from a formal, narrative, poetic, dramatic and psychological point of view, among others.

This is an intuition that is widespread among directors, fans and cinema spectators, and a well-known quote by Francis Ford Coppola sums up this perception of the role of the soundtrack in the overall audiovisual experience: "Sound is the director's best friend because it secretly influences the viewer" (NIETO, 2002: 1).

Indeed, music and image are two structural elements of the audiovisual binomial that touch, encounter, trample, frustrate or even manipulate each other within the general framework of their formal architecture and expressive capacity, although the way in which these relations are established should not necessarily be, as Coppola suggests, secret or mysterious.

It should be possible to establish a theoretical explanation of the role of music and sound within the cinematographic apparatus. In this respect, audiovisual studies of music and soundtracks have proliferated in recent decades, revealing a growing interest in a discipline in which, traditionally, a huge theoretical gap has existed.



Top. Erland Josephson as Domenico. Bottom. Domenico douses himself with petrol after his speech. / Courtesy of Trackmedia

Since the 1990s, the French theorist Michel Chion – researcher, composer of *musique concrète* and intellectual disciple of Pierre Schaeffer – has taken an interest in exploring and attempting to classify this expressive relationship. To Chion’s research we could add the work of other scholars, such as Claudia Gorbman, John Mundy and Kathryn Kalinak, in the list of current relevant research within the discipline, although the studies of these researchers could not always be considered an analysis of music from the perspective of a cinematographic aesthetic, which is the approach taken in this article.

Michel Chion’s contribution to audiovisual studies is a series of innovative perspectives that he developed in his paradigmatic study *Audio-Vision: Sound on Screen* (1994)¹. In essence, Chion postulates that music is a creative and transformative platform of diegetic space-time, the freest element of cinema’s dramatic resources and audiovisual convention, and the discursive element least conditioned by the need for plausibility which, to some extent, characterises every cinematic production. Music *co-irrigates* and *co-structures* the audiovisual discourse (CHION, 1997: 217-220); “music is presence more than medium” (1997: 192).

Along these lines, this article offers a unique case study of the final sequence in Andrei Tarkovsky’s *Nostalgia* (1983), based on the methodological proposal

of the combined analysis of image and soundtrack developed by Michel Chion in *Audio-Vision*.

Tarkovsky’s legacy is one of a unique artist; his body of work is small, but what there is of it has awakened considerable and growing interest. In absolute terms, the soundtrack analysed in this article has a notably poetic and dramatic role in the sequence – following the general tone of the film *Nostalgia*, and being a general constant in Tarkovsky’s filmography – although quantitatively the presence of the music is very limited. “Theoretically,” Tarkovsky wrote, “cinema in its purest form should be able to get along without music” (quoted by Chion, 1997:32).

This context prompts the hypothesis that a few iconic bars of Ludwig Van Beethoven’s *Ninth Symphony* – note the visual adjective used here for a piece of music – are central for structuring a filmic discourse based on Tarkovsky’s particular perspective; in the sequence of the speech of the insane Domenico, played by Erland Josephson, the music recreates, while at the same time synthesising, a humanist discourse that has been building up over the course of the film, and which the chorus of the *Ode to Joy* leads to its dramatic climax.

Sculpting in time and *Ode to Joy* on a radio cassette player

In formal terms, Andrei Tarkovsky’s films are characterised by their revelation of a strong personality that makes them unique, complex and personal works in the context of Western cinema. Moreover, in Tarkovsky’s work lies the seed of an important theoretical component that he elaborates on in his indispensable book *Sculpting in Time: Reflections on the Cinema* (1986): the cinematic aesthetic principle involving man’s development of “the means to *take an impression of time* [...] of reproducing that time on screen as often as he wanted [...]. He acquired a matrix for *actual time*” (TARKOVSKY 1986: 62).

Nostalgia, Tarkovsky’s sixth and penultimate feature film, was the first he shot outside his native USSR. It started out as a Soviet-Italian coproduction, but the Soviet producer Sovin Film quickly withdrew from the project. Tarkovsky’s relationship with the Soviet authorities had become increasingly conflictive since he filmed *Andrei Rublev* (Andrey Rublyov, 1966) and Tarkovsky ultimately opted for personal, political and artistic exile.

The film narrates the feeling of rootlessness suffered by the poet Andrei Gorchakov (performed by the actor Oleg Yankovsky) while he travels through Tuscany tracing the exile steps of Sosnovsky, a late 18th century Russian composer; the context of the work could easily be viewed as an emotional self-portrait of the experience of the film’s director himself at the time.

At the end of the film, the apocalyptic character Domenico gives a heartfelt speech: standing on top of the equestrian statue of Marcus Aurelius in Rome's Capitoline, he offers up a beautiful soliloquy before committing suicide. This is the final sequence, the narrative climax of the film: Domenico performs a humanist monologue, douses himself in petrol, tries a couple of times to light a lighter, and finally sparks the flame that will burn him alive. The scene takes place in parallel to the other decisive action of the plot in which Gorchakoy, following Domenico's previous instructions, crosses the pool of the hot springs, in an excellent sequence shot, holding a lit candle in a physical challenge to a kind of existentialism that the character of Erland Josephson had encouraged in him earlier in the film.

An interesting aspect of the scene under study is that it contains an explicit and direct reference to music. In other words, *music* has a dramatic role in the action being narrated, both in a formal terms (music playing on the screen) and on a dramatic level (the character asks for the music to be played). The music is referenced within the action, *music* as an object of the scene, as companion to a liturgy. "And now, music," says Domenico while he prepares to set himself alight.

This musical element is recreated in the filmic diegesis by means of a mechanical reproduction, playing on a radio cassette player; the music therefore initially has the quality of a diegetic element. But this object within the scene gradually crosses the ethereal barrier into non-filmic (non-diegetic) time, while on the screen we watch Domenico's body engulfed by the flames. Music breaks free of its bounds as a narrated element within the action, as an explicit element, to be transformed into a formal element of audiovisual language (a non-diegetic element) when it fills up the sonic space of the scene. Finally, at the end of the sequence, the music once again becomes an object of the scene within the narrated action; it returns to its initial condition of explicit object, of music playing on a radio cassette player, of diegetic music, after mutating between the different possible filmic spaces of the scene.

This self-referential audiovisual consideration of music, of the role of music in the liturgy developed in the scene (and, by extension, in the drama and the ontology of any film), and the evolution –associated with expressivity and its use within the theoretical apparatus– of the position of music in the sonic space of the sequence, moving alternatively from a diegetic state to a non-diegetic state, makes this sequence an optimal example of the added value of the musical element in the theoretical apparatus of audiovisual language.



Top. The scene is full of silent human figures. Bottom. A character in the scene imitates Domenico's movements. / Courtesy of Trackmedia

An experimental analysis of the sequence from *Nostalghia* according to the analytical method described in *Audio-Vision*

In *Audio-Vision*, Chion describes, develops and theorises a model of analysis that attempts to explore the interaction established between soundtrack and image in the audiovisual framework. The analytical method proposed pursues, according to the author, a threefold objective: firstly, to satisfy pure intellectual curiosity; secondly, to find a theoretical position that will allow a more in-depth analysis of audiovisual structures and aesthetic harmony; and thirdly, to establish an anti-obscurantist exercise in the face of one's own perceptions, one of the theoretical pillars of the concept of *audio-vision*: to identify how one sense influences the perception of others, "What do I see of what I hear? What do I hear of what I see?" (CHION, 1994: 192).

The first part of the combined analysis of music and image is the verbal requirement, a contextual and nominal description of the sequence; to describe the

cinematographic shot and its iconic content and, at the same time, to catalogue the most prominent sonic and musical elements.

The second part of the analysis, the observation procedure, is the one I seek to explore in this article. This is perhaps the most theoretical and analytical element of the method due to its experimental nature, and the one where Chion positions himself most against the grain. This procedure includes the *masking method* and *forced marriage*.

The *masking method* involves separate observation of the elements that make up the filmic discourse, masking the sound objects while observing the visual elements, and vice versa. The goal is to break the audiovisual contract established by the filmmaker and to analyse its main elements separately: to perform a process of dismantling, of deconstruction of the synchronism.

This deconstruction can be put into practice using two procedures. The first dismantling serves to analyse the element of the soundtrack, which will be re-

ferred to here as *acousmatic listening*²; its goal is to analyse the sound object without viewing the sound source from which it emanates. The second dismantling is the procedure that I will refer to as *deaf viewing*,³ in which the sequence is analysed without the soundtrack. Chion explains that the intention of the *masking method* is to give the researcher “the opportunity to hear the sound as it is, and not as the image transforms it and disguises it; it also lets you see the image as it is and not as sound recreates it” (1994: 187).

1. The masking method

Applying *acousmatic listening* to the final sequence of the film *Nostalghia* reveals the presence of a gentle gale suggestive of an open space; a continuous murmur without allowing the listener to distinguish whether it is a breeze or the background noise typical of any recording. This murmur, located deep in the background, contrasts with the foregrounded sound of Domenico’s voice (dubbed by the Italian actor Sergio Fiorentini), whose diction is clear, sweet and rhythmical.

There is a certain evocation of movement. The barking of a dog breaks the crackle of the silence, creating a growing atmospheric tension; at the same time, in the auditory foreground, without reverberation, we hear the successive clicks of a lighter, and a sound that we identify as burning.

An analogue tape sounds faulty, damaged: a technical error is perceived in the attempt to play a piece of music; a sound that proves to be disturbing, phantasmagorical. Suddenly, the tape is fixed and we hear some intelligible music with brass instruments. The spectator probably does not yet recognise the musical fragment playing mechanically in the diegesis. A dog keeps barking violently at a volume very close to that of the music and finally, the score being played can be identified: the famous chorus of Beethoven’s *Ninth Symphony* enters the scene. We no longer hear the dog barking, or the continuous background crackling; the chorus of the *Ninth* has filled the whole sonic space of the sequence.

The sound of the sequence ends when the mechanical failure in the playing of the tape is evident once again; the error is clearly perceived. At once the music recovers the sonic space that it occupied previously, and quickly vanishes. The background crackling is heard again and, from the instant that the music is no longer audible, we hear some dreadful, anguished guttural screams that conclude the sequence.

The analysis of the video track through the exercise of *deaf viewing* reveals that the sequence contains some slow tracking shots and detailed, painterly shot compositions. In this respect, of particular note is the tracking shot passing over the scaffolding that su-

Top. The successive clicks of Domenico’s lighter. Bottom. The heavy traffic and crowds are silenced in the sequence. / Courtesy of Trackmedia



rounds the equestrian statue, ending at Domenico's back; the shot captures a panorama of Rome at sunset, evoking a certain lightness of air, and revealing a massive depth of field, surprising the viewer with the sight of heavy traffic on a ring road in the distance.

The overall sensation as we view this shot is that the internal rhythm of the sequence speeds up: in silence, the camera pans more dynamically and swiftly. Moreover, the images are much noisier than they seemed to be when viewing the scene normally.

It is interesting to note that in the *deaf viewing* of the sequence, the human figures that occupy the scene stand out. These figures are not heard in the original sequence and, when viewing the images without their soundtrack, a much fuller and more tumultuous scene is revealed. The same phenomenon of masking takes place with other elements, such as the sight of a strong wind, quite evident in the silent images, but which, like the crowd, is masked under the soundtrack of the original sequence. Also evident is an acceleration of the action and the editing *in crescendo* as the scene draws to its conclusion. This is expressed in the use of frontal tracking shots that are generally used to gradually close a wide shot.

To sum up, through the *masking method* we find a normally noisy image silenced by the soundtrack, and a piece of music that alternates between the different possible positions within the audiovisual theoretical apparatus.

2. *Forced marriage*

Forced marriage is the next stage in the analysis proposed by Chion as part of the observation procedure. It is an experimental and creative process based on changing the music of the sequence while keeping the same video track. This distorts and alters the original relationship between the audiovisual elements and exposes the random relationship that sometimes exists between the two elements within audiovisual language, opening up a space for creativity in the analysis.

Through this creative manipulation of the sequence we are creating a new synchronism; it is a type of analysis close to the moment of artistic genesis, revealing points of image-music synchronism⁴ that are created in a non-premeditated way, while other original points disappear.

On combining the video track with the new music—a random segment from the *Cantata BWV 54* composed by Johann Sebastian Bach, chosen because it shares a series of attributes with Beethoven's *Ninth* (both are easily identifiable by the spectator as *classical music*), and at the same time, because of the musical characteristics of the piece, Bach proves a good counterpoint to Beethoven, and when viewing this new combination the sequence seems to undergo a temporal expansion;



Domenico sets himself alight. / Courtesy of Trackmedia

the music—of a slower *tempo* than the original—manipulates and dilates the temporal perception of the sequence. The general expressive quality of the sequence is—with a few differentiated features—similar to the original: an aura of certain solemnity in an action that turns into liturgy.

The main difference with the original sequence is that the manipulated version seems to obviate the dramatic conflict of the character, which is so evident and highlighted with the music of Beethoven. In other words, Bach's music seems to transform the drama positively: we expect someone to put out the fire that burns Domenico, that he will not be mortally wounded; the new music creates the sensation that the conflict will be resolved at any moment with no major consequences. With the *Cantata*, the action narrated in the sequence loses its dramatic effect; it seems that it *could not end badly*. This important reflection on how the music determines the drama, the narration, the explicit elements in the sequence of images, is, we believe, one of the main contributions of the methodology and theoretical perspective of Michael Chion.

In the formal aspect of the synchronism, we find that the collapse of the corpse in flames is not a highlighted point of synchrony as it was in the original; highlighted instead is the collapse of a character watching the scene who imitates Domenico's actions from the square. His fall to the ground, moments after the Domenico's, coincides with the resolution of a harmonic progression. This fact turns the point of verticality between the music—harmonic resolution—and image—the character's collapse—into a point of synchronism that did not exist in the original sequence. Thus, a change in the chosen musical segment could vary this point of synchronism, obviating it or making it even more powerful and underlined.



Anguished guttural screams conclude the sequence. / Courtesy of Trackmedia

In general, the manipulated sequence loses dramatic depth and emphasis, although it is worth noting that passing through the different layers of sound of the diegesis results in an expressive effect similar to the one achieved in Tarkovsky's sequence: the abrupt cut in the music due to a mechanical error produces a sensation of emptiness, of a return to reality, of escape from an artifice produced by the *presence* of the music.

We close this analysis with the following technical consideration: both recordings –Beethoven's *Ninth* and Bach's *Cantata*– have a similar frequency range and are therefore located on similar sonic levels. We believe that this may be why Bach's music does not distort the spatial definition or perception of the diegesis (although it does distort the drama) as would probably happen with a more contemporary production with a greater dynamic of frequencies.

Finally, the considerations observed in the observation procedure (including the *masking method* and *forced marriage*) described and applied in this article, enable us to formulate some answers, when we view the sequence again in its original configuration, to the theoretical questions posed in *Audio-vision*: What do I see of what I hear? What do I hear of what I see?

The ultimate intention of the analysis is to consider the artistic and expressive nature, the added value, of music in the cinematographic image, to develop a reflection on what Chion defines as the audiovisual canvas (1994: 212); in other words, a reflection on the mechanisms that the artist uses to take the audiovisual contract, and its language, to the limit of its expressive and dramatic possibilities. These reflections constitute the core of an analysis of the cinematographic aesthetic which, in relation to the sequence from Tarkovsky's film, is expressed in the gradual transition that the music makes through the different layers of sound in the scene.

Conclusion

The method of combined analysis of image and soundtrack proposed by Michel Chion in *Audio-Vision* allows us to explore the relationship established between the two main elements of the audiovisual in the context of its theoretical apparatus, and how this relationship determines the meaning of the filmic discourse in the final sequence of Andrei Tarkovsky's *Nostalgia*. One of the achievements of Chion's analysis is that it aims to be interdisciplinary rather than relying solely on visual or musicological analysis.

In this respect, the analysis of the added value of music has often proven sterile and fruitless, conditioned by the nature and epistemology of music itself, which is sometimes considered to signify nothing. It is our contention that the fruitless results of research in this area have been a consequence of the failure to propose a validated interdisciplinary methodology of analysis of music in the audiovisual framework; in other words, there is a need to establish a method of a combined analysis of image, sound and music, as it has been demonstrated that adopting a methodology based on musicological studies or cinematographic theory leads inevitably to barren ground (FRAILE, 2008: 22). With this in mind, the analytical proposal described in *Audio-Vision* offers a solution to part of this methodological problem.

However, it is important to note that the methodology described by Chion leaves aside elements of analysis that in recent years have taken centre stage in the debate generated around this discipline. One of these would be the concept of the *audiovisualisation* of music (GOODWIN, 1993: 50): the view that the reception of music is more than a musical text, which is in itself an audiovisual text. From this perspective, the sociological, ideological and signifying *preload* of the *Ninth Symphony* would also be analysed, along with how this *preloaded* meaning of the music (MUNDY, 1999: 5) would condition the audiovisual discourse, contributing even more complexity to the interdisciplinary analysis.

Nevertheless, we believe that an analysis based on this methodology demonstrates that the added and expressive value of the use of music in the sequence filmed by Tarkovsky is contingent on the shifting of the music through the different sonic layers of the scene, alternating between the diegetic and non-diegetic space. The reconstruction with Johann Sebastian Bach's *Cantata* manipulates the temporal and dramatic perception of the original scene slightly, as the conflict is obviated and the narrated action loses dramatic force, although the definition of space is not distorted. With the manipulation of the *Cantata*, new points of synthesis appear while others disappear.

The general considerations we observed in the observation procedure are that the soundtrack (tending towards silence) influences the elaboration of the overall sense of the sequence, as it seems that the image is, by itself, prone to movement and noise. A correspondence is observable between the results obtained and the theoretical proposal put forward by Chion, which pursues the objective of identifying sounds in the vacuum (CHION, 1994: 179): the analysis identifies sound objects that are evoked by the image but not mechanically reproduced for viewing, as well as significations that are not represented explicitly through logocentric language, but which emerge from the audiovisual whole through their appearance on the screen.

For future research it would be useful to alternate between diverse exercises of *forced marriage*, using music of different genres and eras, recorded with different techniques –and the latest technology– to analyse how the image would react to these new manipulations and to be able to make a comparative analysis.

In conclusion, this analysis has confirmed that Beethoven's *Ninth Symphony*, contrasted with Bach's music, fits better in the drama of the scene and conditions it; it stretches the audiovisual canvas to the limit of its expressive capacity –reaffirming several of Chion's theoretical propositions, such as the sometimes arbitrary nature of the image-sound relationship– and creates a narrative climax which, in spite of the powerful and distressing images shown, with the chorus of the *Ode to Joy* resembles a lyrical proclamation of humanism characteristic of the cinematic gaze that Andrei Tarkovsky employed throughout his intense career.

Notes

* The images that illustrate this article have been provided voluntarily by the authors of the text. They are frames (screenshots) from *Nostalgia*. We would like to thank Trackmedia for authorizing their reproduction on these pages. (Editor's note).

- 1 Hereinafter this work is referenced as *Audio-Vision*. Michel Chion's academic interest follows a clear line of study sustained in his previous works, such as *Le son au cinéma* (1985), *La musique au cinéma* (1999), and completed in *Music in the Cinema* (1995), or *Film, A Sound Art* (2009), among others.
- 2 The concept of *acousmatic listening* is inherited from Schaeffer's proposition of different types of listening (1998: 159-169), in which the sound source is differentiated from the sound object (1998:49). "The *acousmatic* situation [...] symbolically forbids any relationship with what is visible, touchable or measurable" (1998: 57).
- 3 Based on the differentiation between sound object and sound source, it is considered that images without music or sound may suggest sounds, blows and rhythms. The term *deaf viewing* prompts a reflection on the concept of *silent film* and on the po-

tentuality of the image; the image is not *silent*, yet the spectator –due to a technical incapacity– becomes *deaf*.

- 4 The neologism *synchresis* is defined as the point of synthesis of a synchrony, "the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time [...] independently of any rational logic" (Chion, 1994: 63).

Bibliography

- CHION, Michel (1994). *Audio-Vision: Sound on Screen*. New York: Columbia University Press.
- (1997). *La música en el cine*. Barcelona: Paidós.
- DURAN, Jaume (2012). Los recursos auditivos del cine y el audiovisual. In J. GUSTEMS (ed.), *Música y sonido en los audiovisuales* (pp. 101-114). Barcelona: Publicacions i Edicions de la Universitat de Barcelona.
- FRAILE, Teresa (2008). *La creación musical en el cine español contemporáneo*. Doctoral thesis. Salamanca: Universidad de Salamanca.
- GOULD, Glenn (2001). *Glenn Gould on Television: The Complete CBC Broadcasts, 1954-1977*. [DVD] Toronto: CBC.
- GOODWIN, Andrew (1992). *Dancing in the Distracting Factory: Music Television and Popular Culture*. Minneapolis: University of Minnesota Press.
- MUNDY, John (1999). *Popular Music on Screen: From Hollywood Musical to Music Video*. Manchester: Manchester University Press.
- NIETO, José (2002). *La música para la imagen. La influencia secreta*. Madrid: SGAE.
- SCHAEFFER, Pierre (2006). *Tratado de los objetos musicales*. Madrid: Alianza Editorial.
- TARKOVSKY, Andrei (1986). *Sculpting in Time: Reflections on the Cinema*. Austin: University of Texas Press.

Josep Torelló Oliver (Barcelona, 1982) holds a bachelor's degree in Audiovisual Communication from Universitat de Barcelona. He is currently working on his doctoral thesis, provisionally titled *La música en los films de Pere Portabella (1967-2009)* in the Department of Visual and Plastic Education at the Universitat de Barcelona. He is also a composer and guitarist.

Jaume Duran Castells (Barcelona, 1970) holds a doctorate in Audiovisual Communication from Universitat de Barcelona, a bachelor's degree in Philology and Linguistics and a MAS in Art History. He is a professor at Universitat de Barcelona and collaborates with Enginyeria i Arquitectura La Salle - Universitat Ramon Llull and with the Escola Superior de Cinema i Audiovisuals de Catalunya (ESCAC). He has given national and international courses, seminars and conferences and has a wide range of publications to his credit. He is member of the Board of Directors of the Societat Catalana de Comunicació.