UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
"Scienze-filologico-letterarie, storico-filosofiche e artistiche"

CICLO XXXIII

An Aesthetics of the Audiovisual Gesture: A Neurohumanistic Approach

Coordinatore:
Chiar.mo Prof. Italo Testa

Tutore:
Chiar.mo Prof. Michele Guerra
Chiar.mo Prof. Vittorio Gallese

Dottoranda: Anna Kolesnikov

Anni Accademici 2017-2020

Table of Contents

0. Introduction

In his recent *Film, Art, and the Third Culture: A Naturalized Aesthetics of Film*, Murray Smith (2017) brings Charles Percy Snow's (1956/1998) call for a "third culture" into the heart of the contemporary debate on cinema and the brain. The peculiarity of the current encounter between the sciences and humanities is that it feels both decidedly contemporary and oddly archaic; however, the "anxiety" of the "two cultures divide" actually tends toward the latter, dating back to the turn of the nineteenth century (Snow, 1998: ix). To be sure, distinct domains of human knowledge have existed since antiquity, but during the Middle Ages and Renaissance, in particular, nature was regarded as only a single component in the pursuit of "philosophy". After the achievements of the Scientific Revolution, spearheaded by the discoveries of Isaac Newton, reconsiderations as to what constitutes "genuine knowledge" begin to emerge, leading to the designation of a new discipline – "natural philosophy", i.e., the study of the natural world. Even so, a "continuum" between the study of "human affairs" and the natural world was still in place; for example, the distinguished achievement of the Enlightenment – *L'Encyclopédie* – did not make the distinction between human knowledge obtained through the "sciences" or the "humanities" (Snow, 1998: x). The "fissure", as Stefan Collini notes in his presentation of *The Two Cultures*, began to occur during the Romantic era, from the end of the eighteenth to the beginning of the twentieth centuries, rooted in a fear that calculation and measurement would displace, or replace, the values of "cultivation and compassion", as well as a fear that "secular" knowledge would threaten religious belief (Snow, 1998: xi). Reviving the debate initiated by biologist Thomas Huxley a century earlier, Snow's essay *The New Statesman* (1956) campaigned against what he saw as the crippling division between the traditional "literary intellectual" culture and the culture of the natural sciences, urging for the construction of a "third culture" that would draw upon and integrate the insights of disciplines spanning the arts, humanities, and social and natural sciences (Smith, 2017: 5). An oft-quoted passage of his reads as following:

> "A good many times I have been present at gatherings of people who, by the standards of the traditional culture, are thought highly educated and who have with considerable gusto been expressing their incredulity at the illiteracy of scientists. Once or twice I have been provoked and have asked the company how many of them could describe the Second Law of Thermodynamics. The response

> was cold: it was also negative. Yet I was asking something which is the scientific equivalent of: Have you read a work of Shakespeare's?" (Snow, 1998: 17).

Although he notes that the "two cultures" divide was not as deep in the Soviet Union as in the West (Snow, 1998: 15), it would appear that Snow was unaware of the extent to which his vision of a "third culture" was realized in the years following the October Revolution in 1917 where, as Julia Vassilieva (2019) observes, a complex "testing ground" for a fusion between the arts and sciences was being put into action. Indeed, the Soviet cultural agenda of "reshaping the world" in a "socialist mold" led to intense experimentation in the arts and pioneering research projects in the sciences (Vassilieva, 2019: 24). This unprecedented cultural space that existed until the rise of "Social Realism" under Stalin in the 1930s, in which ideas "migrated" freely across disciplinary divisions, is delineated in Ana Olenina's (2020) newly published book on the relationship between psychophysiology and arts during the Russian avant-garde, as well as in numerous preceding works of scholarship (see: Ivanov, 1976; MacKay, 2007; Tikka, 2008; Bulgakowa, 2014; Toropova, 2017; Widdis, 2017; and Belodubrovskaya, 2018). Nowhere was this "proto-third culture", to borrow Vassilieva's term, more manifest than in the theory and film practice of Sergei Eisenstein – the Soviet polymath in whom the revolutionary ethos of the Soviet avant-garde melded with the Renaissance ideal expressed by Leon Battista Alberti: "a man can do all things if he will".

Eisenstein was an early settler of this "third cultural" continuum, unifying a myriad of disciplines in his investigation of, amongst many other topics, the notion of the "gesture" that pervades all arts and, in particular, cinema, weaving a sensorimotor account of film theory that predated the insights of modern neuroscience. It is for this reason that Eisenstein, in our continuing re-appropriation of his legacy, can be considered as a springboard for "neurohumanistic" research in the "third cultural space". As Naum Kleiman writes:

> "Paradoxically, our image of Eisenstein is also changing all the time. The most positive aspect of this whole process is that he has not yet been canonised. We can still argue about him. Indeed, he does not allow himself to be canonised" (Kleiman, 1993: 32).

An examination of the cultural notion of "gesture" from both a humanistic and neuroscientific point of view, as I intend to do in this thesis, is a tentative to bring Eisenstein's model of third

culture into the twenty-first century. In particular, the debate around "movies and the brain" which initiated over a century ago with Hugo Münsterberg (1916/2004), evolving into the intense season of French filmology in the 1950s (Gastaut and Bert, 1954; Michotte, 1963/2017), and continuing nowadays with cognitive, phenomenological, biocultural and enactivist approaches (Carluccio and Villa, 2006; Konigsberg, 2007; Hasson, 2008; Grodal, 2009; Shimamura, 2013; D'Aloia and Eugeni 2014; Coëgnarts and Kravanja, 2015; Carocci, 2018; Eugeni, 2018), provides fertile grounds for neuro-humanist research. In particular, the discovery of mirror neurons and the direct link between perception and action (de Pellegrino et al., 1992; Gallese et al., 1996), the resultant theory of embodied simulation (Gallese, 2005) and the interdisciplinary research it has fueled (Freedberg and Gallese, 2007; Gallese and Guerra, 2020) stand not only to "revitalize mimesis", as Gallese has noted, but also to provide the necessary empirical framework with which to operationalize Eisenstein's initial triangulation of the theories of objective psychology and psychophysiology (specifically those of William James (1884) and Vladimir Bekhterev (1913)) with film aesthetics.

Specifically, the following questions will be addressed in this thesis:

1. *In light of Eisenstein's writings, what is the role that he attributes to the role of the body in shaping the notion of the actor's, musical and audiovisual gesture in film?*
2. *How does Eisenstein's body of work on the notion of gesture relate to those of other thinkers, both historical and contemporary?*
3. *How do Eisenstein's ideas on gesture and audiovisual correspondence relate to contemporary debates in neuroscience?*

In light of these research questions, a "triangulated" approach will be employed in order to address the following considerations:

1. *What can Eisenstein's ideas on the notion of gesture contribute to understanding of the cognitive and neural underpinnings of audiovisual experience in film?*
2. *How can empirical testing of conceptual notions of gesture, such as Eisenstein's, inform film theory?*

3. *How can the creation of "naturalistic stimuli" through collaboration with the arts contribute to experimental methods and advance understanding of the role of corporeal engagement in shaping audio/visual experience?*

The notion of "third culture" contends the dire need for a methodological reform enabling the "constructive integration" of knowledge from the natural sciences with those of the humanities without displacement or replacement of one discipline's conceptual framework with another (Vassilieva, 2019: 24). In *Film, Art, and Third Culture*, Smith describes what Barbara M. Stafford calls an "equally knowledgeable Janus-faced method" (Stafford, 2011: 4) or, in the context of film studies, an approach that:

> "[…] while fully acknowledging the diversity of artistic forms and their cultural contexts, sees film art as a manifestation of a cluster of deeply entrenched, basic human capacities, and thus treats it as a phenomenon which is likely to be illuminated by various types of scientific as well as traditional humanistic research" (Smith, 2017: 4).

Within such a framework, Smith elaborates, "the questions posed determine what methods are brought into play, without the traditional barriers […] intervening (Smith, 2017: 4). Specifically, this can be achieved through triangulation, which involves "fixing" the object in explanatory space by "projecting lines from each body of evidence, and following them to see where they intersect" (Smith, 2017: 60). Going further, Smith explains that:

> "what distinguishes triangulation as a form of consilience is that it occurs across distinct types of evidence and levels of enquiry, rather than meshing together different bodies of evidence on a single level—as one might, say, in seeking to integrate neural accounts of the perceptual and emotion systems; or psychological theories of memory and imagination. And in triangulating the phenomenological, the psychological, and the neurophysiological levels, our explanations are thickened" (Smith, 2017: 68).

Given the neurohumanistic approach adopted in this thesis, the text at hand will be arranged in the following manner: a "constellation" of the notion of the gesture as explored in Eisenstein's writings as well as in the ideas of antecedent, contemporary and subsequent

scholarship will be discussed in Chapters 1 – 3, an empirical investigation of the theoretical notion of gesture as grounded in Eisenstein's body of work using the methods of cognitive science in Chapters 5 – 7, and as a link between the "two cultures", the "creative" project of stimuli construction through collaboration with the arts will be detailed in Chapter 4.

Specifically, in Chapter 1, I shall discuss the formative and lasting influence of Vsevolod Meyerhold's principles of biomechanics, in themselves shaped by theories of rhythm, empathy, and objective psychology, on Eisenstein's ideas on the role of the actor's expressive movement in facilitating "mimetic contagion" with the stage and film audience, and how these ideas relate to early film and gesture theory. In Chapter 2, I will explore how Eisenstein's collaboration with Sergei Prokofiev on *Aleksandr Nevskii* (1938) and his 1940 staging of Richard Wagner's *Die Walküre* (The Valkyrie, 1876) influenced his ideas on gesture as the unifying link between sound and image, and on the work of art as a "totality", as well as critique of Wagner's influence on the "culture industry" and Eisenstein's ideas on audiovisual unity by Theodor W. Adorno, Max Horkheimer and Hanns Eisler. In Chapter 3, I shall consider Eisenstein's writings on rhythm, pathos and organic unity and the choreographed pathos of his final film *Ivan Groznyi Chast' II* (Ivan the Terrible Part II, 1958), as well as the occurrence of "syncretic" and "synthetic" gestures of "rotation" and "rising" beyond the cinema of Eisenstein.

In Chapter 4, I shall present an "aesthetic model" for the construction of the custom-made, naturalistic stimuli set – the "staircase" scene from Mikhail Kalatozov's *Letiat Zhuravli* (The Cranes Are Flying, 1957), itself an embodiment of the gestures of rotation and rising – and the creative collaborations that were involved in filming the video clips and composing the musical tracks.

In Chapters 5 and 6, I shall report the materials, methods and results of the stimuli validations, as well as discussions of the results with respect to the scientific literature reviews presented at the beginning of the chapters. Specifically, in Chapter 5 an empirical investigation of the effect of Camera Movement, Actor Presence and Image speed on participant ratings will be presented, and in Chapter 6, the impact of Composition Contour, Composition Complexity and Note Pattern on participant ratings will be reported. In Chapter 7, in light of the validated and selected stimuli, I will present the materials and methods of the audiovisual experiments,

i.e., the behavioral and electroencephalography (EEG) studies. Due to the restrictions of the Covid-19 pandemic, EEG data collection was postponed, and only behavioral results will be reported and discussed. Specifically, an investigation of the effect of Modality (i.e., Audio, Video, or Audiovisual) and Contour (i.e., Ascending or Still) on participants ratings will be reported.

Finally, I will present a summary of Conclusions, with a discussion of what the empirical results can tell us about the cultural notion of gesture, and I will end with a consideration of directions for future research.

1. The actor's gesture

Eisenstein's formative years as an aspiring theatre and subsequently film director were influenced by two masters of the Soviet avant-garde: Lev Kuleshov and Vsevolod Meyerhold. Kuleshov, largely considered one of the most inspiring figures of montage theory, writes that his three-month long mentorship of Eisenstein during his involvement with the "Collective" was grounded in what would later become known as the "Kuleshov effect" (Kolesnikov, 2020h: 67), and Eisenstein's last living moments were spent on an unfinished letter to him about color film (Bulgakowa, 2002: 48). Whereas Kuleshov's notion of montage as something that is built "brick by brick" consolidated Eisenstein's perspective on the structural essence of film (as Mikhail Iampolski (1991) notes, Kuleshov's theory of montage was also rooted in an interest in rhythm and corporeality), it was Meyerhold who awakened in Eisenstein a lifelong captivation with the power of the arts to "seize" the beholder on a primordial, visceral level (see: Law and Gordon, 1996; Cervini, 2008; Grespi, 2019). Years later, Eisenstein recalled that seeing Meyerhold's production of Henrik Ibsen's "Dollhouse" was an artistic epiphany, writing: "I remember the continuous chills. I was not cold; it was excitement, it was nerves strained to their limits" (Law and Gordon, 1996: 20).

In this section, I will discuss how, through the entirely novel "montage" and application of the ideas of Meyerhold, James, Bekhterev, Rudolf Bode, Jean d'Udine, Émile Jacque-Dalcroze, and Theodor Lipps to theater aesthetics and subsequently cinema, 1) Eisenstein addresses the theme of *mimetic contagion* in film theory, offering a sensorimotor account of how expressive movement on stage and on screen can facilitate the transmission of emotions between the actor and the spectator; 2) how and why such a line of research has been finding solid application in the contemporary neurocognitive debate on film and how it matches embodied simulation theory; and 3) the parallels that can be drawn between Eisenstein's ideas on expressive movement and early film and *gesture theory*, namely the ideas of Béla Balázs, Walter Benjamin and Giorgio Agamben.

1.1 Meyerhold and biomechanics

At the helm of the Constructivist movement, Meyerhold transformed the aesthetic program of Soviet theatre in the years leading up to the Revolution and foremost afterward Fig. 1.1 –

1.5). As Irina Schulzki notes (2019), the theatre had incorporated the teachings of François Delsarte on acting style and Émile Jacque-Dalcroze's method of eurythmics, which became a fixation in the avant-garde scene largely through its propagator, Prince Sergei Volkonsky. Although one of his most brilliant students, Meyerhold rejected Konstantin Stanislavsky's doctrine of psychological fusion with the character: Meyerhold reasoned that it was futile, as well as needlessly trying for the actor, to go through the trouble of *perezhivanie* (experiencing, or introspectively reliving a memory), in order to better inhabit a role, if the spectator sitting in the auditorium has no way of connecting with the actor's emotional process. What can be considered as Meyerhold's manifesto for a *visceral* route to the spectator reads as follows:

> "No pauses, no psychologizing, no "experiencing", either on stage or when preparing for a role [...] Lots of light, joy, grandeur and infectiousness, light creativity, and the public's participation in the actions and collective creation of the performance – this is our theatre program" (Meyerhold, 1921: 16, my translation).

At the heart of Meyerhold's approach to actor training lies: 1) the prioritization of "stylizing" as a fundamental principle of all arts, 2) the recognition of the stage and the auditorium as part of a single organic whole and the audience as an active participant in the work of theatre (in opposition to Stanislavsky's principle of the fourth wall), and 3) the crucial importance of movement (Law and Gordon, 1996: 23). Meyerhold writes:

> "Movement is the most powerful mode of expression in the creation of a theatre production. Deprived of words, costumes, the limelight, the aisles, and the auditorium, and left only with the actor and his mastery of movement, the theatre would still be the theatre. The spectator must come to learn the thoughts and motivations of the actor through his movement, his gestures and his facial expressions" (Meyerhold, 1914: 82, my translation).

In order to achieve the aforementioned objective, the actor's body had to be trained to fluency in "expressive vocabulary", ready to assume a pose, leap onto a chair, or roll off the stage in the most fluid and efficient way possible at the command of the director. A set of 16 biomechanical *études* (exercises) were adopted from theatre traditions such as the *commedia*

dell'arte and Elizebethean and Asian theatre in order to train the actor to gain full and graceful command of his or her body. Trained to develop a musical ear, the études were designed to help the actors develop a sense of rhythm to their movements, pausing when necessary, speeding up and slowing down when the emotional tone called for it. Piano music was used to guide the actor in setting an emotional "tempo" and in developing the skill of punctuating movement with the rhythmic flow of music. Ultimately, Meyerhold's program pushed the actors to strive for the "loose" and "relaxed" but highly controlled movements of expert dancers, what he refered to as *tantsost'* (danceness) (Meyerhold, 1914: 94). As we will see later in this section, the ideal of the Meyerholdian actor closely relates to the aesthetic program of the interwar German *Tantztheater* (dance theater) and new forms of *Ausdruckstanz* (expressionist dance) developed by the pioneering choreographers Mary Wigman, Kurt Jooss and Rudolf Laban (Ruprecht, 2015). However, as Lucia Ruprecht notes, the new gestural dance was not intended for "direct access" to emotions (as delineated by the ethos of biomechanics), but rather served to evoke a "split" between emotion and expression and engender gestural "abstraction" reflective of universal "laws of movement"(Ruprecht, 2015: 23).

In his externalization of the "internal", Meyerhold liberated the actor from the earlier isolation of the stage, encouraging a flow of communication that was at once gestural and semiotic. A former student, identified as "V. Es.", writes:

> "Each movement is a hieroglyphe with its own meaning. On the stage, there can only be movements that can be instantly decoded, otherwise they are superfluous [...] We are not concerned with the physic world, but the physical world [...] This does not mean, of course, that the actor becomes an automaton without a mind. No. Within the acrobat-actor, there is always a duality" (V. Es., cited in Law and Gordon, 1996: 142).

The "hieroglyphe" body of the Meyerholdian "acrobat-actor" is both encoded and "decoded" through gestural meaning, and the conditioning of this corporeal semiology occurs through the development of the "conditioned reflex" at the mark of the director, in turn leading to the emotional "excitation" of the actor as a result of a given movement or posture. It is in the mastery of movement and one's body that the art of the actor resides, and through the firm

principles of conditioning the actor and the audience become a single, organic whole, a "physiological chain reaction" of emotions, with movement serving as both the trigger and the medium.

Basilov, another former student, writes:

> "The peculiarity of the actor's art [...] is that the artistic process takes place in front of an audience, as a result of which the actor and the spectator find themselves in a kind of relationship, which places the spectator in the position of *a living resonator* responding to all the manifestiations of the actor's profession (emphasis added). Conversely [...] the actor reacts instantly through improvisation to all demands that arise from it [the auditorium]. By way of various signs (noise, movement, cough, laughter and others), the actor must identify with precision the spectator's attitude toward the production" (Basilov, cited in Law and Gordon, 1996: 153).

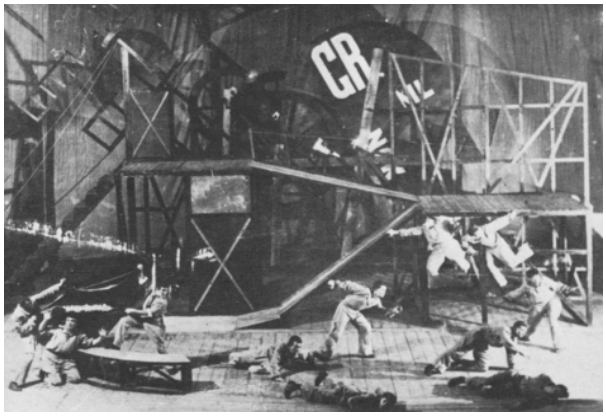
Such notions as "conditioning", "reflex", and "resonance" can be traced to objective psychology theory, notably the work of Bekhterev (1913) and James (1884), as well as to the philosopher Lipps (1906), as we will see later. Meyerhold's profound interest in the orbit of psychology and philosophy of his time had a formative influence on his approach to training actors and making theatre art, and, notably, on the filmmaking and film theory of Eisenstein.

1.2 Eisenstein and expressive movement

In 1923, Meyerhold appointed Eisenstein, his loyal and prodigious student, together with playwright Sergei Tretyakov, to author an article entitled "Vyrazitel'noe dvizhenie" (Expressive movement; Eisenstein, 1964) that would explain his teachings on biomechanics. As many of Eisenstein's writings, it was unpublished at the time, integrated into his 1924 essay "Montazh kino attraktsionov" (Montage of film attractions, see: Eisenstein, 2010a), which was also unpublished at the time, finally finding its way into a chapter on expressive movement in his final opus *Metod* (Method; Eisenstein, 2002) (Neuberger, 2017: 260). In addition to Delsarte and Dalcroze (studied by everyone in the theatrical avant-garde), Ludwig Klages' and Rudolf Bode's text *Ausdrucksgymnastik* (Expressive Gymnastics, 1922) in particular served as

1.1

1.2



1.3

1.4

1.5

Fig. 1.1 – 1.5 | Meyerhold's theatre and biomechanics. Meyerhold's 1928 production of the *Le cocu magnifique* (The Magnificent Cuckhold, 1921) by Fernand Crommelynk, source: Worrall, 1973 (1.1, 1.3); Meyerhold and Erast Garin rehearsing for Nikolai Gogol's *Revizor* (Inspector General, 1836) in 1926, source: Law and Gordon, 1996 (1.2); Meyerhold's 1930 production of Vladimir Mayakovsky's *Bania* (The Bathhouse, 1929), source: *80 let*, 2020 (1.4); workshop, source: *Vsevolod Meierhold,* 2016 (1.5).

foundations for Eisenstein's conceptualization and extension of the corporeal semiology of biomechanics.

In their essay, Eisenstein and Tretyakov translate Bode's text and discuss his analysis of the French neurologist Guillaume Duchenne's research on movement, with Duchenne's thesis reverberating even in *Metod*: "Le movement musculaire isolé n'existe pas dans la nature" (Isolated muscular movement does not exist in nature) (Eisenstein, 2002: 147). Bode conceptualized Duchenne's teachings into the principle of *Totalität* (totality; *tselekupnost'* (Russian)), achieved through expressive gymnastics and "organic rhythm" training, which was translated into an abundance of exercises, tables and scientific terminology, all of which Eisenstein brought to the Proletkult theatre. By performing a motor process with "organic correctness", as when in a natural state, a condition of "affective conception" and "plasticity" is achieved. Going beyond Meyerhold's and Bode's teachings, Eisenstein called this the "attractive potential of movement". Specifically, an "attraction" is an emphasized movement on stage that serves to evoke "infectious emotional interaction" – i.e., *mimetic contagion*. According to the actor Alexander Levshin, it was during the opening night of Eisenstein's 1923 production of Alexander Ostrovsky's *Na vsyakogo mudretsa dovolno prostoty* (Enough Stupidity in Every Wise Man, 1868) that his theory of the "montage of attractions" (1923) (see: Eisenstein, 2010a) took full form (Levshin, cited in Law and Gordon, 1996: 172). In his inquiry of whether it is possible to measure "the activity of art", Eisenstein begins to give contour to what would become the determinate position of his political-artistic philosophy – the director's vision must bypass individualities in the spectators through the visceral forging of their bodily and mental states (Belodubrovskaya, 2018). Eisenstein writes:

> "The attractiveness of an expressive movement [...] is assured [...] to the extent that expressive movement ensures the excitement of the emotions intended for the viewer. It is precisely expressive movement, built on a correct biological foundation, that is solely capable of arousing this emotion in the spectator, who in turn repeats *reflexively*, in a *weakened* form, the whole system of the actor's movements. Following this production of movements, the spectator's *incipient muscular tensions* are released in the desired emotion" (Eisenstein, 1964: 154, my translation and emphasis).

Indeed, with his notion of "incipient muscular tensions", Eisenstein came to the same theoretical conclusion that Vittorio Gallese and colleagues would come to empirically decades later with the discovery of mirror neurons in macaques and the theory of embodied simulation that followed – we tacitly simulate the movements we observe in others (Tikka, 2008). It would become part of Proletkult folklore that Eisenstein, emulating his mentor, would sit with his back to the stage, scrutinizing every bodily expression and facial expression in the audience, occasionally rising up to "induce a bout of tears or an armful of laughter" or even a scare that would "force them to jump out of their seats in horror" (Eisenstein, 1964: 139).

In addition to Bode, Delsarte and Dalcroze, a forgotten theoretical source that had a formative influence on Eisenstein is *L'art et le geste* (Art and Gesture, 1910) by music theoretician Jean d'Udine, translated into Russian shortly after its publication by Prince Volkonsky (Olenina and Schulzki, 2017; Neuberger, 2017). This largely overlooked study on the origins of art would play a key influence on Eisenstein's idea of audiovisual synesthesia and the gesture as the "embryo" of expressive form and all art, a notion which will be examined in closer detail in later sections. Indeed, D'Udine sees "creative gestures" in practically all the arts (Schulzki, 2019: 106). The movement, rhythms and gestures of nature must be "repeated" and "crystallized" in the artwork in the process of exteriorizing the artist's emotions. In response to this, the beholder "vibrates" with the assimilated impressions (as we will see in Chapter 3, Eisenstein elaborates the notion of rhythm in his writings on the rhythmic drum and organic unity). One of the concepts that Eisenstein built on in his framework is that of *rakurs* (taken from the French "raccourci"), signifying a "foreshortening", or visual transformation of an object's shape or angle into an unusual position for the viewer. Eisenstein writes that *rakurs* is an "important, fixed moment" during a production – a technique that permits the "crystallization of essential gestures" to be imprinted in the memory of the spectator as "a series of photographs" (Law and Gordon, 1996: 98). It serves a "utilitarian" purpose as a "point of rupture between two movements", enabling the actor to summon the required "excitation". In contrast, *poza* (pose), Eisenstein explains, is always static and "self-contained", completed *a priori* for the purpose of "pleasant observation" (Eisenstein, 2002: 68). In one of Eisenstein's last texts, "Mise en Jeu" (Bringing into play, 1948; i.e., the manifestation of the filmmaker's vision into the actions dictated by plot) and "Mise en Geste" (Putting into action, 1948; i.e.,

the gestures and movements characters perform), Eisenstein returns to d'Udine's formula, writing that the director must render the characters' inner conflicts "tangible" for the audience through movement, actions and gestures (Eisenstein, 2004). It is notable that the American film theorist Victor Oscar Freeburg (1918), in a very different cultural context, wrote of corresponding "pictorial moments" and, in particular, "tonal" and "linear rhythm" (Freeburg, 1918: 26). As Neuberger notes, although Eisenstein's formulations may appear "naïve", "simplistic" or "impractical", it is precisely these methods that Eisenstein used in his masterpiece *Ivan Groznyi* (Ivan the Terrible, filmed in 1944 – 1946, released in 1958) with "unforgettable results" (Neuberger, 2017: 263).

In *Ivan*, Eisenstein explores the 16th century Tsar's (played by Nikolay Cherkasov) gradual descent into madness. Interweaving Ivan's reflections - doubt for his chosen path, remorse for his actions - Eisenstein paints a psychological portrait of "tragic duality" and "physical anxiety". The lonelier Ivan becomes, the more he pursues his goal of unifying Russia. The "expressive" staging is decisive in conveying the themes of the film, such as betrayal, terror, paranoia and loneliness. As Neuberger explains, Ivan's transformation from "vulnerable child" to "bloodthirsty tyrant" is incarnated in the "double dialectical" conflict within him, meant to evoke "pathos" and "synthetic moments of transformation" (Neuberger, 2017: 269). The task that Eisenstein attributed to Cherkasov was to visibly convey this inner conflict through movement, gesture, and facial expressions. During production, Eisenstein studied Meyerhold's personal archives (which he agreed to safeguard after his arrest in 1939) and finally discovered the "methodological secrets" that his old mentor had declined to reveal. According to Naum Kleiman, *Ivan* is a prologue to the opus on expressive movement that Eisenstein had wanted to write. According to cast members, Eisenstein had long discussions with his actors about the "intellectual and emotional profile" of their characters, but all "psychological meaning" was evoked through "physical pose and bodily movement" (Neuberger, 2017: 259). The expressiveness in *Ivan* is largely constructed on *otkaznoe dvizhenie* (recoil movement), one of the pillars of biomechanics. When looking at something, Cherkasov would crane his neck forward in a "bird-like fashion" before suddenly stepping back (Bordwell, 2005: 240). Yuri Norstein, the Soviet and Russian animator, argues that the acting in *Ivan* is actually modelled on animation, where "persona" is expressed by "distilling" it into one particular gesture or "look," and one that is often "zoological" (Neuberger, 2017: 267). Cherkasov famously complained of all the "contorted" positions and movements that

Eisenstein required of him, however he also credited Eisenstein with significantly enriching and broadening his "technical ability" as an actor and his sense of movement in space (Neuberger, 2017: 256).

Indeed, it is the notion of *otkaznoe dvizhenie* (Fig. 1.6 – 1.8) that Eisenstein examined with the most scrutiny in his conceptualization of expressive movement. According to Eisenstein, it is one of the "fundamental laws" of expressive movement – rather than going from point *A* to point *B* directly, it is necessary to first withdraw to point *C*, which is situated in the opposite direction to point *B,* and only then can one initiate the movement toward point *B* (i.e., A – C – B). If one needs to jump up on a ledge, one first recoils backward in the opposite direction in order to increase the inertia of motion and thus gain momentum. If a tree needs to be chopped, one first swings the axe in the opposite direction before coming down with full force. And when throwing a punch, similarly, the arm and upper body first draw back from point B to point C before delivering the blow to point A (Eisenstein, 1964: 84). Eisenstein refers to this phenomenon as an "organic law" that is applied "always" and "everywhere" in everyday life, and in the theatre it serves to deliver "an additional blow to the spectator's psyche". It abides by the law of the "negation of the negation", in which the movement from C to A cancels the initial negation of moving from A to C, rather than A to B directly, and the final A to B acquires a "qualitatively different meaning". Furthermore, this phenomenon exists not only in physical phenomena but also psychic processes: a doubt is simply another form of recoiling (C – A) before taking a decision (C – A– B).

Later, in *Metod*, Eisenstein returns to the subject of the "primacy" and "organic qualities" of movement across all life forms, noting that even primitive unicellular organisms such as amoeba exhibit continuous contractions that are "correctly" repeated, what Eisenstein refers to as "vegetative movement" (Vogman, 2020). Eisenstein asserts, "all biological functions have a rhythm, except consciousness and will" – the more primitive the organism, the more basic the rhythm (Eisenstein, 2002: 194). The basic rule that one should never go directly "from intention to desire", as the duration must always be "extended" in order to transform an act into a "process", can be seen in the "medicinal pipe ceremony" of the Native American Sioux tribe. During this ceremony, the shaman stands up and slowly dances in a circle until he reaches the sacred pipes. He stops and bows as if to open the package, but as soon as his hand is about to touch it, he suddenly draws back, and this "false movement" is customarily

repeated three or four times before finally touching the sacred object (Eisenstein, 2002: 204). According to Eisenstein, this negation of negation, or denying of the denial, enables the actor to access the spectator's perception, opening the path to the spectator's emotions. The recoil (or refusal), through dialectical and physical contrast, both enhances the transmission of emotions between actor and spectator and creates *otsraenie* (i.e., distancing, or defamiliarization), heightening the perception of the familiar and encouraging the spectator to recognize the aesthetic language of the theatre.

Toward the end of the nineteenth century, a new movement developed in Imperial Russia and abroad, termed "objective psychology", that was firmly founded on the outright rejection of the "soul" and a general disinterest in the "unconscious" – the central interest was that of measuring behavioral processes. Two key figures, James and Bekhterev, in parallel made key contributions to the understanding of human physiology and stand out in their tremendous influence on the cultural project of Constructivism.

A philosopher, medical doctor and psychologist, James' radical empiricism significantly altered modern psychology's understanding of the physiological mechanisms of emotions, notably canonized in the "James-Lange" theory of emotions. James' (and independently Carl Lange's) hypothesis contributed to a somatic re-interpretation of emotions as something that is first felt physiologically and only afterward apprehended cognitively, commonly recapitulated as: we are sad because we cry. If one were to extract the characteristic "bodily symptoms" of a strong emotion, all that would be left is a "cold", neutral state of cognition. A "purely disembodied human emotion", James declares, "is a non-entity" (James, 1884: 194).

Transient states of emotional awareness are directly mediated by the body, and it's the physiological response to a given stimulus (i.e., a bear, an insult, a loss) that is the emotion itself – not the mental awareness of the change in bodily state. James' theory had a direct influence on Eisenstein's writings on expressive movement, opening his eyes to the potential of theater (and art in general) to generate visceral reactions in the beholder. Eisenstein writes, "James' view of emotion [...] as a psychologically subjective state derived from a physiological approach here finds full application". For Eisenstein, James' theory is decisive for the audience – the spectator reproduces the movements seen on stage, and enters the emotional state of the actor. Eisenstein writes, "the physical movement of a man gives rise to a particular

1.6

1.7

1.8

Fig. 1.6 – 1.8 | ***Otkaznoe dvizhenie* (recoil)**. Recoil in the "slap" (1.6-1-7), source: Law and Gordon, 1996; example of biomechanical exercise entitled "archery", source: Iablonska, 2014 (1.8).

excitation for that movement. Thus, laughter evokes a good mood, tears - sadness. William James" (Eisenstein, 1964: 136).

In parallel to James, the neurologist Vladimir Bekhterev established the science of objective psychology (i.e., "reflexology") in Imperial Russia. All behavior, according to Bekhterev, can be explained by the empirical study of reflexes, and internal mental states certainly exist and are of great import, but are unamenable to the scientific method. Bekhterev divided the mechanisms of mimetic movements into two categories: "reflexive" and "associative". Whereas reflex mimetic movements, such as laughter when tickled, occur automatically in response to a direct stimulus, associative mimetic movements, such as an expression of suffering upon hearing distressing news, is a response to an "indirect" stimulus. Tickling can directly produce laughter, but hearing a sound does not usually cause pain unless it is *associated* as such, in which case it can cause as much difficulty in the retention of tears as a physical injury (Bekhterev, 1913: 300). The association, according to Bekhterev, emerges as a result of an individual's previous experience. Whereas James is cited directly in Eisenstein's writings, direct references to Bekhterev are far fewer, often leading to the false assumption that it was perhaps Ivan Pavlov's work on conditioning that influenced Eisenstein's theory of the montage of attractions. However, Eisenstein was not familiar with Pavlov's work before 1925 – it was through Bekhterev's research that Meyerhold and then Eisenstein incorporated objective psychology into their theory of the transmission of emotions between the actor and the audience (Bordwell, 2005: 89). Although the Nobel-prize winning Ivan Pavlov is considered the father of classical conditioning, what Bekhterev refers to as the "associated reflex" is in fact analogous to Pavlov's "conditioned reflex". Pavlov's rise to worldwide fame was due to the American psychologist John B. Watson, who discovered Pavlov's salivation research and incorporated it into his theory of behaviorism. In contrast, Bekhterev's essential contributions to psychology and neurology were largely ignored. However, Bekhterev's notions of reflexive mimetic movements (innate physiological responses) and associated mimetic movements (conditioned by society or previous incidents) are apparent in Meyerhold's (and later Eisenstein's) theories on acting. Indeed, the new Soviet theatre audience, after attending a few Meyerhold productions (and later a few Eisenstein films), would begin to associate certain emotional responses with certain expressive movements.

In addition to James and Bekhterev's objective psychology theory, Eisenstein also assimilated Lipps' (1906) philosophical contributions to the notion of *Einfühlung* (feeling into), i.e, how people experience both aesthetic objects and the mental states of others. For Lipps, *Einfühlung* is a process of "inner imitation" based on a "natural instinct" that makes us imitate the movements and expressions that we perceive in others. We experience the feelings of the other as our own because we project our own feelings onto the other, and by way of this projection, we enter into "union" with the other. According to Lipps, we simulate the entire repertoire of movements when attending a ballet performance, but we do not *act* – the body is not itself activated, and aesthetic experience is instead psychologized. Indeed, having been familiar with his theories ("Thus, it is not a question of the 'sincerity' of the movement of the actor, but of his imitative and mimetic fidelity {see Lipps}" (Eisenstein, 1964: 151)), Eisenstein applied Lipps' ideas to this theoretical and practical explorations of psychological empathy in cinema. At first glance, it may seem that Eisenstein, like Lipps, saw empathy or *Einfühlung* as a "fusion" of self and other. Through motor resonance, we project ourselves into the Lippsian acrobat or the Meyerholdian actor-acrobat, and his or her actions become our actions and our emotions. However, Eisenstein's psychomotor account of acting in the theatre and cinema does not align with Lipps' translocation of the self into the other, as James' physiological account of emotions and Bekhterev's theory of the associated reflex are inextricably linked to the sensorimotor capabilities of the body. Indeed, in Eisenstein's lecture at La Sarraz conference on Independent Film ("Imitation as Mastery", 1929; see: Eisenstein, 2006), he asserts:

> "The actor was and is the most direct object of imitation. We never know precisely *what* is going on inside another person. We see his expression. We mimic it. We empathise. And we draw the conclusion that he must feel the same way as we feel at that moment. The actor shows us how to feel" (Eisenstein, 2006: 15).

Eisenstein rejects the dissolution of the self in Lippsian empathy, adopting a perspective that relates to that of the proponents of the phenomenological school of empathy such as Edith Stein and Max Scheler, who emphasize the necessity of attributing to the other, rather than the self, that which is emotionally at stake. The Soviet director's view on empathy also corresponds to Robert Vischer's notion that the physiological dimension of aesthetic empathy

is vital, and that in our empathy with moving objects, we animate our own "immobility", "going up with ascending lines" (Vischer, 1873, cited in Pinotti, 2016: 160).

There is an increasing realization that Eisenstein's account of visceral spectatorship is very much aligned with the contemporary theory of embodied simulation based on the discovery of mirror neurons (Tikka, 2008; Belodubrovskaya, 2018; Vassilieva, 2019; Olenina, 2020; Gallese and Guerra, 2020). We understand the meaning of the other's gesture because our nervous system reacts *as if* it was us doing it. The Shared Manifold Hypothesis (Gallese, 2001), which describes a system of "self-reflection" that continuously models the other onto self and seeks to explain how the activation of the mirror neuron system in the human brain occurs, manifesting at the phenomenological level as empathy, is absolutely consistent with Meyerhold's and Eisenstein's idea of expressive movement as a vehicle for mimetic contagion on the stage and across the screen. Although Eisenstein was influenced by Lipps' ideas on motor imitation and resonance, the basis of his ideas on the transmission of emotions between the actor's body and the body of the spectator fits into the theories of James and Bekhterev (which assert that emotions are specific physiological states), thus ruling out a full Lippsian projection of the self onto the other. In going further than perhaps any other film theorist and practitioner in his assimilation of psychomotor theories in psychology and aesthetics, Eisenstein's film theory is particularly amenable to empirical research (as will be addressed directly later in this thesis).

1.3 Early film and gesture theory

As Tsivian (2008) observes, early cinema was an "oscillation" between the "opposing" elements of expression, montage and gesture which together engender all meaning in film. Indeed, the turn of the twentieth century marked a "kinaesthetic" turn in the European cultural sphere, a time of unprecedented interest in movement across the arts and sciences. Defined as "feeling through movement", *kinaesthesia* is the act of seeking, through muscle cognition, a felt understanding of subjective experience (Sirotkina, 2014, 2019). Coincidentally, it culminated in parallel to the birth of the cinematograph – the synthesis of artistic expression and industrial feat, and the ultimate movement machine. It's predecessor, chronotography, was invented by one of France's leading physiologists, Etienne Jules Marey, for the scientific

study of movement and the exposure of what had previously remained invisible to the human eye (Fig. 1.8 – 1.9).

Starting in the 1860s until the turn of the twentieth century, Marey analyzed and froze bodies in motion using complex optical devices that captured moving figures as a rapid succession of images, photographed against the backdrop of an "artificial darkness" (Elcott, 2016). In his use of artificial darkness for the composition of his music-drama theatres, Wagner (whom we will discuss in Chapter 2) famously proclaimed that "time here [chronotography] becomes space" (Wagner, in Elcott, 2016: 5). It was above all the human figure, "moving and creating space," that enabled cinema to become what it was – the relation between moment and movement (Didi-Huberman, 2020). Dominated by "self-conscious" displays of "exaggerated gesticulation", the first decade of film captured a wide array of athletic activities such as dancing, juggling, tumbling, fencing and boxing. This early exhibitionism is canonized in the adoption of Eisenstein's notion of "cinematic attraction" to denote the exhibitionism and privileging of spectacle over narrative in the first decade of cinema, in particular (Gunning, 1986; Gaudreault and Gunning, 2006). As Jonathan Auerbach writes in *Body Shots*, "early filmmaking revealed the rhetoric of the human form, turning the body itself into an expressive medium" (Auerbach, 2007: 17).

As Gallese and Michele Guerra note (2020), this primacy of movement and "Action!" was recognized by early film theorists such as Freeburg, Henry Albert Phillips and Epes Winthrop Sargent, during a time in which the United States saw a surge in literature on filmmaking and "photoplay writing" in the years between 1912 and the early 1920s. This period in film theory would serve to "swiftly stabilize" classical American style in which the various meanings attributed to action accomplish narrative purposes. In *Technique of the Photoplay* (1916), Sargent defines action as "(a) Any gesture performed by the player" and "(b) The various actions of individual players whereby the narrative is advanced" (Sargent, 1916: 161). A year later, Phillips in *The Photodrama* (1914) writes:

> "The perfect photoplay leaves no doubts, offers no explanations, offers nothing it cannot finish. It is all action, action, ACTION! And by action we mean technically visualized interpretation of whatsoever nature that convincingly

> contributes to the perfect illusion of emotionally seeing a dramatic story (Phillips, 1914: 64).

However, Phillips emphasizes that action "in and of itself", as clearly manifested in the "chase movie" is not sufficient – action has to be incorporated into the "dramaturgical structure" of the film. Phillips' formula reverberates with the ethos of Meyerhold's biomechanical theatre, where there must be "no pauses, psychologizing or 'experiencing'", the spectator must be able to infer the character's motivations and emotions (i.e., psychological profile), from the embodied meaning of their movements, gestures and facial expressions, where movement and mind are not estranged but rather an "organic whole". A few years later, in *The Art of Photoplay Making* (1918), Freeburg pairs aesthetic motion and the expression of the human body (as in dance) to physiological impact, writing:

> "The reaction of our senses to the human form and physical movement is primary and fundamental, and takes place before our brain has time to interpret the dramatic meaning of the visual stimulus" (Freeburg, 1918: 12).

Like Eisenstein, Freeburg anticipates embodied simulation theory and echoes James' somatic theory of emotions and 19th century notions of *Einfühlung*, calling into question our impulse to "project our very selves on the screen" at the sight of "mere shadows on the screen". Going further, Freeburg writes:

> "the sense of an incipient or impending stir outside us stimulates a corresponding agitation within us, and we are impressed before we really know what it is that impressed us" (Freeburg, 1918: 12).

In 1920, two French physicians, Edouard Toulouse and Raoul Mourgue, presented the results of an investigation of the impact of various film genres on spectators' respiration at the Congrès de L'Association Française pour l'Avancement des Science in Strasbourg, concluding that the "motor suggestibility of the participants" is what drives film's "very strong influence on feelings and emotions" (Toulouse and Mourgue, 1920, cited in Gallese and Guerra, 2020: 77). As Jean Epstein would later write in *Le Cinéma du Diable* (The Cinema of

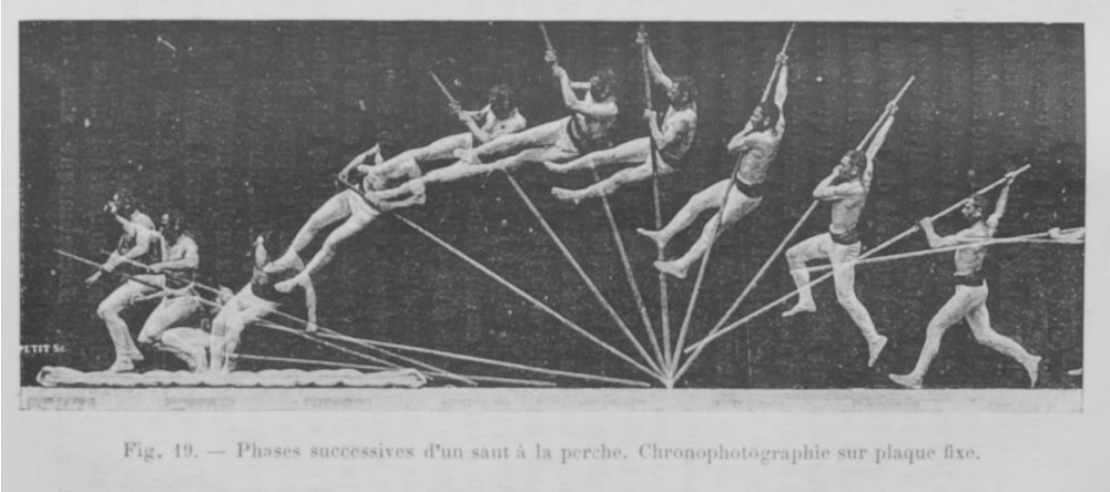
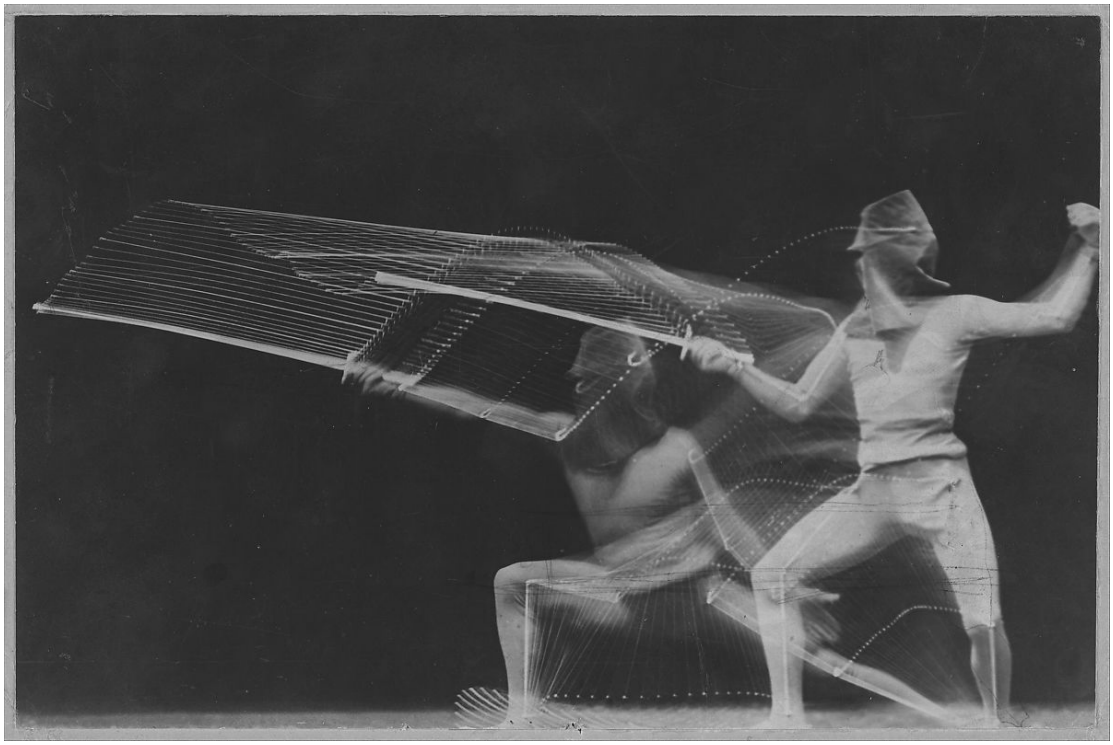


Fig. 19. — Phases successives d'un saut à la perche. Chronophotographie sur plaque fixe.

1.8



1.9

Fig. 1.8 – 1.9 | **Deconstructing the perfect gesture**. Chronotographes of athletes pole vaulting (1892), source: MIT libraries, 2011 (1.8); and fencing (1906) (1.9), source: The Metropolitan Museum of Art.

the Devil, 1947), "the cinematic spectacle foremost brings into play older, and thus more fundamental, faculties which are qualified as primitive: those of being moved" (Epstein, 1947: 22).

As Andrea Pinotti and Antonio Somaini discuss in their *Cultura visuale. Immagini, sguardi, media, dispositivi* (Visual culture. Images, gazes, media, devices, 2016), the birth of cinema signaled the "re-ascendancy of immediacy" (*unmittelbar*), particularly for Béla Balázs. In the first of his three volumes on cinema, *Der sichtbare Mensch* (*Visible Man*, 1924/2010), Balázs extols the arrival of a new *visuelle Kultur* (visual culture) that would at long last liberate man from the restrictions of verbal language, a legacy of the printing press, returning him instead to the primacy of images and gestures. Balázs writes:

> "The whole of mankind is now busy relearning the long-forgotten language of gestures and facial expressions. This language is not the substitute for words characteristic of the sign language of the death and dumb, but the visual corollary of human souls immediately made flesh. *Man will become visible once again*" (Balázs, 2010: 10).

In this new realm, man would express himself not with words but with the "language of mimicry and gestures" – both the new and "aboriginal" mother tongue of the human race (Pinotti and Somaini, 2016: 4). The close-up, which would canonize the face of Renée Falconetti in Carl Theodor Dreyer's *Jeanne d'Arc* (1927), is recognized by Balázs as the unique invention of cinema and a preeminent vehicle for *Mienenlehre* (new gestural language) which, from a Marxist standpoint, would serve to topple social and national barriers and create an "international human type". Balázs asserts that the "variety of facial expressions and bodily gestures has drawn more frontiers between people than has any customs barrier, but these will gradually be eroded by film because the film market operates on the principle of 'one universal language of gesture'" (Balázs, 2010: 14). As Abe Geil notes, Balázs' gesture is primarily expressive and is "purified" by film of its "historical relation to language" (Geil, 2018: 516). On the one hand, film is a means of "anamnesis" of "lost language gestures and facial expressions", and on the other hand it is a "machine" for "standardizing" gesture and expression, driven by the profit principle of the market (Geil, 2018: 515). A few years later in *der Geist des Films*

(Spirit of Film, 1930/2012), Balázs writes, "If we look at and understand each other's faces and gestures, we not only understand, we also learn to feel each other's emotions" (Balázs, 2012: 44). However, as Geil observes, underlying Balázs utopian vision of the advent of "global monolingualism" and the messianic metaphor of "loss and redemption" runs an idealization of the "uniform type of the white race" and the "unique, shared psyche of the white man", although Balázs scholars point out that in *der Geist des Films* this "racial essentialism" is eventually abandoned (Geil, 2018: 514). Siegfried Kracauer, for example, critiques Balázs argument that through film the immediacy of the "international human type" would overpower capitalism, claiming that this only holds on to "the bad rationality of capitalist thinking" and that only through the "insights that come with language can radical change take place" (Kracauer, 1927 cited in Geil, 2018: 515; see also: Kracauer, 1995).

In his essay "Notes on Gesture" (1992/1993), Giorgio Agamben discusses the late 19th century crisis that befell European culture. Agamben writes, "In the cinema, a society that has lost its gestures seeks to re-appropriate what it has lost while simultaneously recording that loss" (Agamben, 1993: 137). Referring to physicians such as Gilles de la Tourette, Agamben discusses the proliferation of nervous pathologies and motor disorders, a symptom of the "loss of control" over gesture in the wake of a destabilized social order, and a "historical unhinging" of gesture from subjectivity (we, the heirs of modernity, live in a "generalized Tourettism") (Geil, 2018: 521). In parallel, a society that has lost its gestures is destined to "gestural infatuation". Agamben offers various examples of this fixation, most notably Aby Warburg's unfinished atlas "Mnemosyne", which the Italian philosopher describes as "not a fixed repertoire of images, but virtually a moving representation of the gestures of Western humankind from classical Greece to Fascism [...] (Agamben, 1993: 138). Indeed, Friedrich Nietzsche's *Also sprach Zarathustra* (Thus Spoke Zarathustra, 1883/2005) is "the ballet of a humanity bereft of its gestures", and his concept of "eternal return", as well as Sigmund Freud's "unconscious", are attempts to remedy the loss of subjectivity in modernity as exhibited by cinema (Baumbach, 2019: 180). According to Agamben, gesture, rather than the image, is the true element of the cinema, and for this reason it "ranks with ethics and politics". However, as Agamben notes in "L'autore come gesto" (The Author as Gesture, 2005/2007), it is "neither production nor enactment" – it is suspended between expression and lack of

expression, "what remains unexpressed in each expressive act" (Agamben, 2007: 66). As many of Agamben's notions on the gesture, they originate in Benjamin's media theory.

Somaini writes, "What we may consider as Benjamin's 'media theory' is [...] the study of the interplay between the historically-evolving domain of the technical and material *Apparat*, and the '*Medium* of perception' in which they operate, determining different 'articulations of the real'" (Somaini, 2016: 10; Pinotti and Somaini, 2012). In his seminal essay "Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkei" (The Work of Art in the Age of its Technological Reproducibility, 1935/2008), Benjamin affirms that the human perceptual field of modernity would become sidetracked by technology, introducing his notion of "innervation". Within this historical context, Benjamin envisiones that the new mode of mass perception will be assembled through people's collective "innervation" to technologies. The screen actor is the primary vehicle of this process, able to "spark *collective* innervation at the level of *reception*, in the corporeal space of the audience assembled in the theatre", not through a direct relationship with the audience, but with the "apparatus", which leads to a "recomposition of the actor's being" (Ruprecht, 2015; Hansen, 2004). Benjamin was influenced by Bertolt Brecht's notion of *Gestus*, which denotes "the realm of attitudes adopted by the characters toward one another", including "physical attitude, tone of voice and facial expression". When these actions take on an "allegorical function", it becomes "social *Gestus*", and the ability to "typify social behavior" and render it visible is what Benjamin refers to as its "quotability". In his 1934 essay "Franz Kafka. Zur zehnten Wiederkehr seines Todestages" (Franz Kafka: On the Tenth Anniversary of His Death), Benjamin recognizes the "uncanniness" of gesture as a state that is simultaneously concrete and intangible, proposing that, "The invention of motion pictures and the phonograph came in an age of maximum alienation of men from one another, of unpredictably intervening relationships which have become their only ones" (Benjamin, 1999: 814). In another essay entitled "Was ist das epische Theater?" (What is the Epic Theater?, 1931), Benjamin writes that, "Epic theatre is gestural [...] the more frequently we interrupt someone engaged in an action, the more gestures we obtain. The interruption of an action is thus at the foreground of epic theater" (Benjamin, 1998: 3). With regard to theatre, Benjamin offers an account of gestural "arrest" as a defamiliarizing theatrical intervention that deconstructs everyday gestural "regimes" (Ruprecht, 2015: 28). Gesture in the epic theatre is conceived neither as a "universal language of the soul" nor an

"involuntary physiological response" (Ruprecht, 2015: 23). The condition that epic theatre reveals is the "dialectic at a standstill", born not out of contradiction between words or behavior, but "of the gesture itself". Returning to cinema and Benjamin's notion of collective innervation as "gestural force" (closely linked with his idea of "mimetic faculty" and its migration into the aesthetic affordances of visual media in the "optic unconscious"), Miriam Hansen observes that it may have less in common with Freudian psychoanalysis than with the objective and perceptual psychology and kinaesthetic theories of art of the time, in particular the discourse of biomechanics that must have reached Benjamin through Asja Lācis. Hansen writes:

> "[…] what seems important regarding Benjamin's concept of innervation and its implications for film theory is the notion of a physiologically "contagious" or "infectious" movement that would trigger emotional affects in the viewer, a form of mimetic identification […] The recourse to neurophysiological and reflex psychology may not be as sophisticated as the insights of psychoanalysis; yet it may have been more in tune with new, technologically mediated forms of aesthetic experience, predicated on mass production, unprecedented circulation and mobility, and collective, public reception" (Hansen, 2011: 137).

Indeed, evoking biomechanics in "Was ist das epische Theater?", Benjamin writes:

> "The Russian producer Meyerhold was recently asked in Berlin what in his opinion distinguished his actors from those of Western Europe. He replied: 'Two things. First, they think; second, they do not think idealistically but materialistically.' The view that the stage is a moral institution is justified only in relation to a theatre that does not merely transmit knowledge but actually engenders it. In epic theatre the actor's training consists in acting in such a way that he is oriented towards knowledge; and this knowledge, in turn, determines not only the content but also the tempi, pauses and stresses of his whole performance. This should not, however, be understood in the sense of a style." (Benjamin, 1998: 10)

And in "Der Sürrealismus. Die letzte Momentaufnahme der europäischen Intelligenz" (Surrealism: The Last Snapshot of the European Intellegentsia, 1929), Benjamin asserts:

> "Only when in technology body and image space so interpenetrate that all revolutionary tension becomes bodily collective innervation, and all the bodily innervations of the collective become revolutionary discharge, has reality transcended itself to the extent demanded by the Communist Manifesto" (Benjamin, 2005: 217).

In evoking his notion of innervation, Benjamin describes what is precisely Eisenstein's achievement in the mimetically powerful "Vakoulintchouk's funeral" scene of *Bronenosets Potemkin* (The Battleship Potemkin, 1925). The crowd's sentiment shifts from grief to rebellious excitement as the revolutionary energy of the clenched fist rapidly infects the crowd, calling to action and solidarity. Indeed, Benjamin's notion of collective innervation to technology and Eisenstein's visceral spectatorship share a common influence in the theories of objective psychology.

* * *

In the next section, an analysis of Eisenstein's ideas on the expressive gesture will be extended from stage and screen acting to the film as an embodied whole, in particular Eisenstein's ideas on movement as link to audiovisual correspondence in the sound film, and the influence of Wagner and d'Udine on his theory of vertical montage.

2. The musical gesture

In one of his lectures in 1946-1947, Eisenstein told his film students the following:

> "I have to confess that I am not even very educated, musically speaking. But fate decreed that I should work with two wonderful musicians, one who belongs to the past and the other indissolubly joined to contemporary culture. With the first I have had a deep association through his works and with the second through direct collaboration. Working with both has taught me a great deal and helped me with a lot of what I had intuitively felt with regard to music (and composition generally) as forms taking shape in a visually unfolding process. One of them is Wagner and the other Prokofiev [...]" (Eisenstein, 1968-74, cited in 1992: 62).

In this section, I will explore the Prokofiev and Wagner "collaborations" and their influence on the Soviet director, and the primacy of the gesture audiovisual that permeated Eisenstein's mature writings on audiovisual unity. Specifically, I will discuss, 1) Eisenstein's theory of vertical montage and embodiment of emotional theme as analyzed in *Aleksandr Nevskii* (1938); 2) the notions of *Gesamtkunstwerk* (total work of art) and *leitmotif* in Wagner's *Der Ring des Nibelungen* (The Ring of the Nibelung, or The Ring Cycle, 1876), and Eisenstein's interpretation of these ideas in his production of *Die Walküre* (The Valkyrie); and 3) Adorno's, Horkheimer's and Eisler's critique of the "culture industry" as well as the Wagnerian/Eisensteian approach to audiovisual unity.

2.1 Prokofiev and vertical montage

Like in *Ivan* six years later, the most relevant and inspiring creative aspect of *Nevskii* (1938) was Eisenstein's collaboration with the celebrated composer Sergei Prokofiev. Following the Russian Revolution, Prokofiev permanently settled in Moscow in 1936 after a long sojourn abroad, during which Prokofiev had worked with the likes of Sergei Diaghilev on his *Ballets Russes* in Paris, such as *Le pas d'acier* (The Leap of Steel, Stal'noi skok, 1926) and *Le Fils prodigue* (The Prodigal Son, Bludnyi syn, 1929), the latter choreographed by George Balanchine. In the Soviet Union, Prokofiev began to simplify his creative style, such as for one of his most famous works, *Petia i volk* (Peter and the Wolf, 1936), for Natalia Sats' Central Children's Theatre, a "symphonic fairy tale for children" which was subsequently adapted

almost a decade later by Walt Disney as a renowned animated short in 1946. During this period Prokofiev also wrote several choral pieces, the first Soviet opera *Semen Kotko* (1940) and a cantata for the twentieth anniversary of the October Revolution, continuing the genre of "song-symphony" that became widespread in the 1930s. From the Eisenstein-Prokofiev correspondence (1939-1946) assembled and translated by Ronald Levaco, it is clear that the two artists shared a mutual veneration for each other. Indeed, when approached to write the music for a film in 1948, Prokofiev responded that "Since the death of Sergei Eisenstein, I consider my motion picture activities terminated" (Levaco, 1973: 2). Prokofiev, too, left a mark on Eisenstein. In particular, the film director was impressed by Prokofiev's ability to quickly capture the emotional effect, rhythm and structure of a scene, "create images in music" and, upon viewing a scene, provide a fully scored "musical equivalent of the visual representation" by the next day. Eisenstein considered Prokofiev the "perfect composer for the screen", whose music was "amazingly plastic", but as Levaco notes, Eisenstein's films were analogously symphonic, conceived in elliptical structure around a central theme that ran through the film and served as a "unifying leitmotif" (Levaco, 1973: 4).

According to the Soviet director, art begins when the combination of sound and image do not simply reproduce a relatedness, "but establishes one". In *Nevskii,* Eisenstein and Prokofiev would take turns in providing the creative "template" – sometimes the scenes would be edited in accordance with Prokofiev's music pre-recorded on the soundtrack, and other times the music was entirely written to fit a fully completed visual montage. In order to achieve what Eisenstein refers to as *sinhronost'* (synchronicity) between image and sound, the filmmaker must be sensitive to the "fundamental themes and motifs" that underly his or her work, and the key to this "audiovisual commensurability" is movement. Eisenstein recalls that he would project the rushes for Prokofiev, often in rough cut, explaining in "words, sounds and gestures" what he desired musically for the scene.

Nevskii (1938) recounts a key event in Russian history: Prince Alexander Nevsky's opposition to the invasion of the Teutonic Knights, in particular the battle of Lake Peipus which in 1242 ended their eastern expansion. *Nevskii* came after Eisenstein's "humiliating confession" for the "mistakes" of his suppressed *Bezhin lug* (Bezhin Medow, 1937), in which the director declared that the subject of his new film would only be "heroic in spirit, militant in content, and popular in its style" (Bordwell, 2005: 220). Providing a clear appeal for patriotic loyalty to

Russians fearing an imminent German invasion, *Nevskii* is considered Eisenstein's most accessible work, finally securing Eisenstein an Order of Lenin and worldwide success.

In his essay "Montazh 1937" (Montage 1937; see: Eisenstein, 1964), Eisenstein puts forward an account of melodic music as *gestural* and closely linked to human physiology. The voice, with its tones and modulations, for example, results from the physical vibration of the vocal chords which, due to their high frequency, is perceptible only as sounds to our hearing. The voice, Eisenstein posits, is the "forerunner" of melody and the "source" from which we draw the "basic human emotional content" in a musical melody (Eisenstein, 1964: 253). As we will see in Chapter 6, this finds strong application in Arnie Cox's "human voice as source domain" theory on embodied music cognition. Rhythm, in turn, originates in the primitive movements of work and dance, which themselves are the forerunners of "music proper", as we will see in the next chapter. Eisenstein expands his gestural account of music in "Vertikal'nyi montazh" (Vertical Montage, 1940), where he asserts that when listening to music anyone can "visualize certain graphic images", and we can also depict "with a movement of our hand, the impression a certain detail of a piece of music makes on us" (Eisenstein, 2010b: 374). Eisenstein's perspective on music relates to the aesthetics of the classical world, where Plato posited that music reproduces and imitates human emotions, and in this lies the source of its emotional power. Similar to his teacher, Aristotle claims that the emotional power of music lies in its ability to imitate emotions. Despite the difficulty in attaining total congruence between pictorial and musical elements due to an inherently subjective nature that resists "any methodological scheme", Eisenstein asserts that any "concrete" correspondence between image and sound still lies in the *gesture* (Eisenstein, 2010b: 373). He writes:

> "For music — the "gesture" which underlies both sound and picture — is not something abstract and unrelated to the theme, but is the most generalised expressive embodiment of the image through which the theme is articulated" (Eisenstein, 2010b: 376).

Similarly, in "Chetvertoe izmerenie v kino" (The Fourth Dimension of Cinema, 1929; see: Eisenstein, 2010a), Eisenstein asserts that:

> "[…] for musical overtones, the term 'I hear' is no longer strictly appropriate. Nor 'I see' for the visual. For both, we introduce a new formula: 'I feel'" (Eisenstein, 2010a: 186).

For Eisenstein, the notion of gesture translates from the actor's body to the film as an audiovisual whole, and it is in the linear trace of movement that the audiovisual link is to be constructed. In the progression of a melody, there is a kinesthetic element that shapes visual and embodied forms during music perception, and this gesture must "embody the emotional theme" of a shot or a sequence in in its entirety. Eisenstein writes:

> "[…] in matters strictly confined to shot composition, too, the human being and the interaction between his *gesture* and *intonation* when they are generated by a single emotion, will prove to be the determining prototype of shaping audiovisual structures which […] derive from a single determining image" (Eisenstein, 2010b: 328).

As we will see, the primacy of the gesture fits into the Soviet director's greater aesthetic project in which the ultimate goal is to "demolish all contradictions" between the visible and audible worlds, creating a "unity", a *synesthesia,* that everyone from the Greeks to Alexander Scriabin has dreamed of (Eisenstein, 2010b: 337). To this end, *Nevskii* served as a testing ground for Eisenstein's developing theory on the interrelation between movement, the human body, and the creation of "maximal audiovisual correspondence".

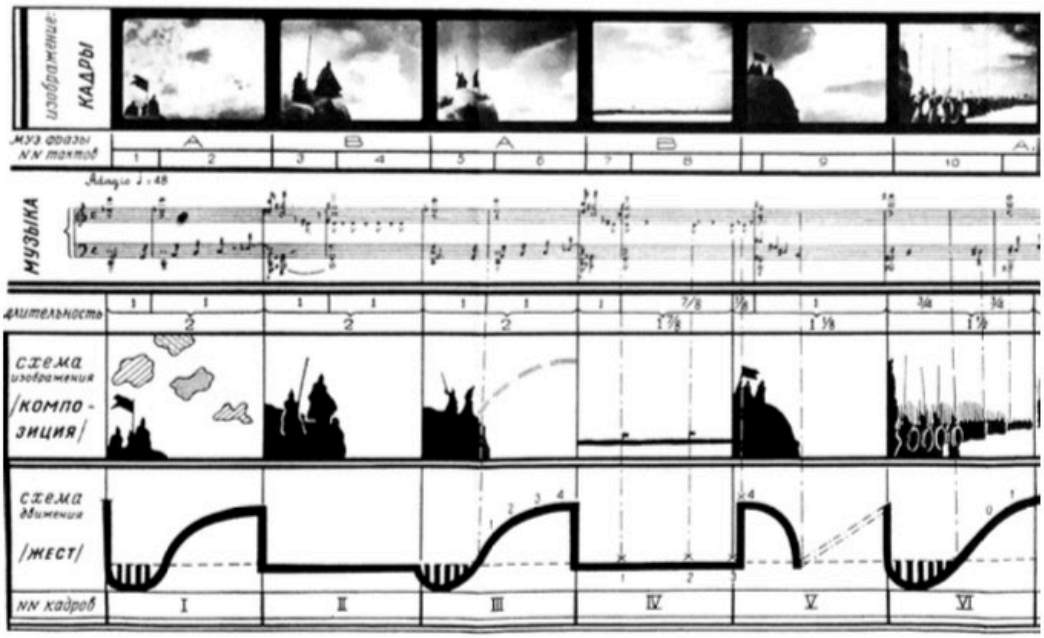
Eisenstein regarded the eponymous "Battle on Ice" sequence as the film's "furthest reach of experimentation", and indeed was the first sequence of film to be shot, in the summer of 1938 (Bordwell, 2005: 221). In "Vertikaln'yi montazh", Eisenstein presents his well-known *a posteriori* analysis of the sequence, unveiling the "secret of vertical correspondences" which link the music "step-by-step" across the pictorial composition of the first twelve, static shots of the sequence (i.e., vertical montage) (Fig. 2.1 – 2.2). Eisenstein adopts the term "vertical" from the visual of an orchestral score (as we will see later, a visual that was strongly evoked during Eisenstein's study of Wagner for his production of *Die Walküre*), adapting it to the idea of an "audiovisual score" in which a vertical line passing from one frame to the next symbolizes the sequential progression of a film sequence and graphically corresponds to the

"melodic gesture" of the music (Fig. 2.3). Serving as a kind of "physiognomy" of an audiovisual sequence (and an uncanny predecessor to the use "eye-tracking" in psychology), it illustrates Eisenstein's hypothesis on the attentional gaze of the viewer, in which the movement of the eye over the lines of the pictorial composition of each shot, coinciding with the melodic gesture of the music, is not simply related to vision, but an embodied act of perception. As our eye moves along the outline of the static, pictorial features of the shot, Eisenstein predicts, we "transfer" the "outline" of the movement that our eyes make to the shot itself, evoking an embodied act of perception that corresponds to the kinesthetic experience evoked by the melodic gesture. Eisenstein writes that there must be:

> "[...] *congruence* between a *musical* sequence and a *pictorial* sequence, which would enable one to *combine vertically*, i.e., *simultaneously*, the progression of each phrase of music in parallel with the progression of the [...] shots; and this must be done within conditions that must be as strictly observed as those within which we *combine pictures 'horizontally', i.e., sequentially*" (Eisenstein, 2010b: 370).

In his *a posteriori* analysis of *Nevskii* and vertical montage, Eisenstein also discusses his theory of "embodiment of emotional theme". Eisenstein explains that the thematic emotional content of the "Battle on Ice" sequence is embodied in the "rhythm of the anxious expectation" of impending battle that permeates both the pictorial and musical composition. Nevsky's army awaits the charge of the Teutonic Knights, gazing silently into the distance over the frozen Lake Peipus. Even before the battle is fought, without physical contact between the armies, the approach of the impending hostile forces can already be "felt". For example, in the transition between the third and fourth shots (Fig. 2.1), the melodic phrase expresses the act of "holding one's breath, when the chest is ready to burst from the rising tension inseparable from the forced inhalation" (Eisenstein, 2010b: 391). When the emotional tension approaches its "peak", it is followed by a "sudden release, a total decrease in strain, like the release of a "repressed sigh".

When the Teutonic Knights begin their charge across the frozen lake (Fig. 2.4), Eisenstein asserts that the "intense throbbing of the heart" dictated the "rhythm of the galloping hooves" (Eisenstein, 2010b: 392). He also notes that in addition to the tonal, linear, special and graphic elements that establish "inner synchronicity" between image and sound, it is the



2.1

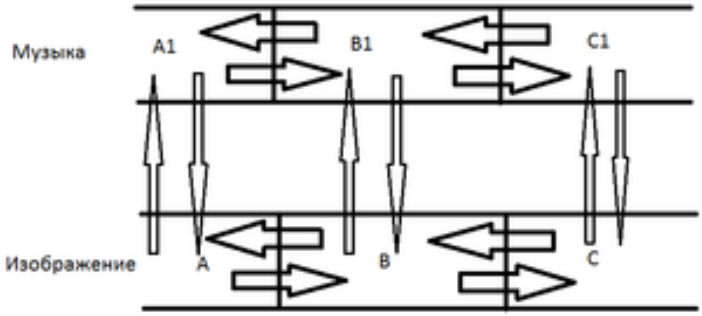
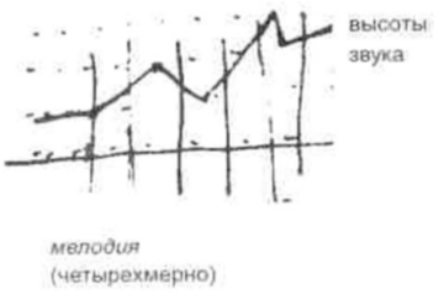


Схема 2

2.2

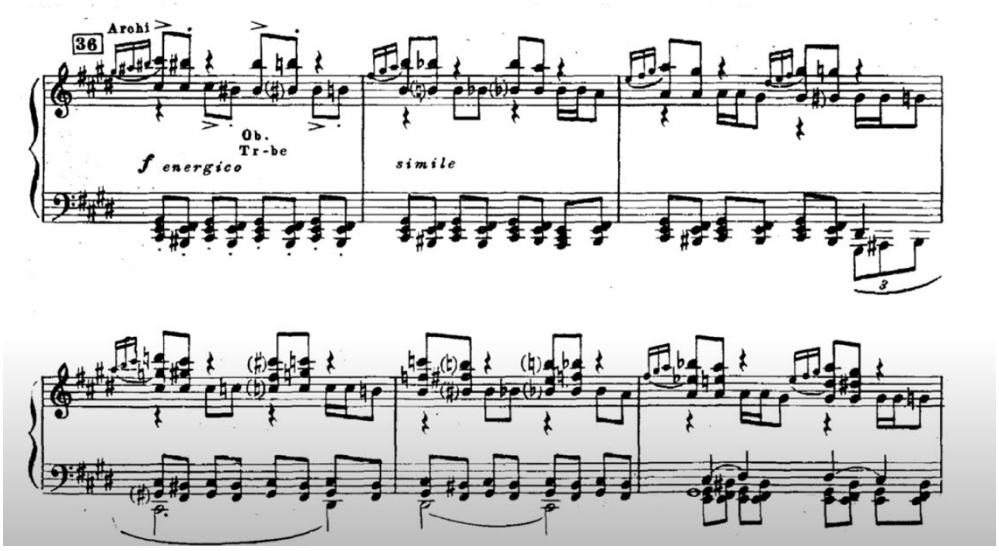


2.3a

2.3b

Fig. 2.1 – 2.3 | Vertical montage in *Aleksandr Nevskii* (Eisenstein, 1938). Diagram attached to Eisenstein's essay "Vertikal'nyi montazh" (Vertical Montage, 1940) depicting the first 12 still shots of the "Battle Over Lake Peipus" sequence (*izobrazhenie kadry*: picture frames) and the piano version of Prokofiev's orchestral partition of the sequence (*muzyka*, music) (2.1); diagram depicting the relation between music (*muzyka*) and image (*izobrazhenie*) in "Diagram 2" (*Shema 2*) from "Melodia i zhest" (Melody and Gesture, 1947) (2.2); diagram depicting a four-dimensional melody (*melodia-chetyrehmerno*) and its pitch height (*vysota zvuka*) from diagrams from "Melodia i zhest" (2.3a); and a three-dimensional drawing (*risunok (trehmerno)*) from "Melodia i zhest"; (2.3b). Source: Eisenstein, 2004.

2.4



2.5

Fig. 2.4 – 2.5 | Embodiment of emotional theme in *Aleksandr Nevskii* (Eisenstein, 1938). Still from the "Battle on Ice - Charge of the Teutonic Knights" sequence (2.4); Prokofiev's musical score "Aleksandr Nevskii Cantata, Op. 78: "Battle on Ice" (2.5). Source: *Prokofiev,* 2011.

"psychological, acting factor" that drives the "rising intensity of emotion" and is equally important for the "embodiment" of gesture (Eisenstein, 2010b: 392). The "Battle on Ice" sequence, specifically, is meant to evoke "fear and dread" of imminent death. Indeed, Prokofiev's ominous Cantata Op.78 (see Fig. 2.5) bears a particular similarity to Carl Orff's frightening *Carmina Burana* (1935), foregrounded visually by the nightmarish appearance of the Teutonic Knights' armor. Prokofiev was initially disheartened by the film's poor sound-mixing, but soon discovered that certain technical difficulties could be exploited, and made use of the harsh quality of the sound to heighten the chilling effect of the bass motifs (Levaco, 1973: 2). Later, Prokofiev would adapt his score into a large-scale Cantata for mezzo-soprano, orchestra and chorus and was widely performed.

Eisenstein's approach to achieving bodily empathic engagement through the construction of audiovisual correspondence "by feel and touch" resonates with the contemporary framework of embodied cinema, as we will see in later chapters. In order to shed light on the origin of Eisenstein's ideas on audiovisual unity, the Soviet director's "indirect" collaboration with Wagner (i.e., his production of *Die Walküre*) while writing "Vertikalny'i montazh" in 1940 will be considered.

2.2 Wagner and the embodiment of myth

As Martin Puchner notes, Wagner conceived a theory of gesture that forged a synthetic link between drama, music and poetry, influencing theatre, performance, literature, and in particular, cinema (Puchner, 1999: 26). Wagner's new artform, termed the *Musikdrama* (music drama), transcended the Romantic traditions of musical performance to create "multimodal events" striving to achieve what Wagner termed *Gesamtkunstwerk*, i.e., the total work of art. Following the failure of the 1848 revolutions, the German composer wrote in "Eine Mitteilung an meine Freunde" (A Communication to My Friends, 1851):

> "I shall never write an *Opera* more. As I have no wish to invent an arbitrary title for my works, I will call them Dramas [...] I propose to produce my myth in three complete dramas, preceded by a lengthy Prelude (Vorspiel) [...] At a specially appointed Festival, I propose, some future time, to produce those three Dramas with their Prelude, *in the course of three days and a fore-evening*. The object of this

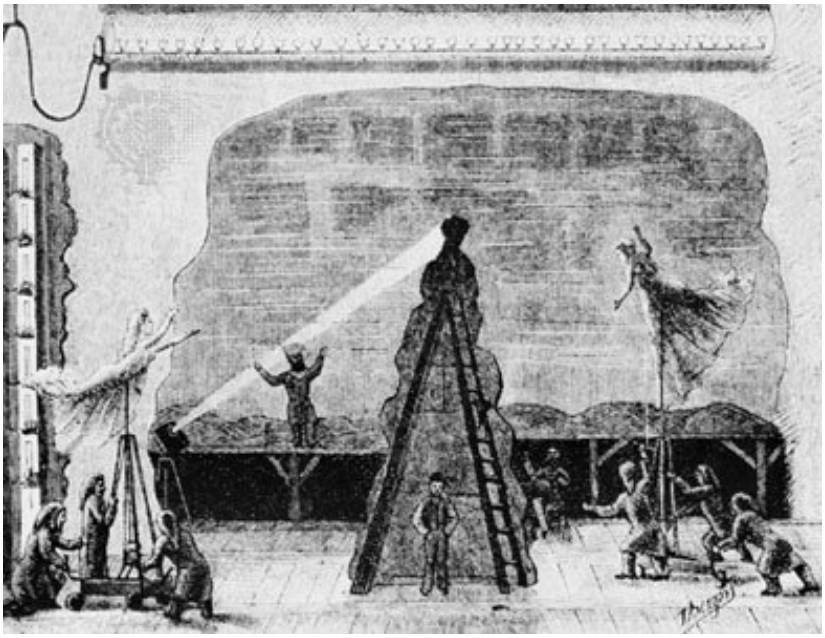
> production I shall consider thoroughly attained, if I and my artistic comrades, the actual performers, shall within these four evenings succeed in *artistically conveying my purpose to the true Emotional* (not the Critical) *Understanding* of spectators who shall have gathered together expressly to learn it" (Wagner, 1993: 351).

A cycle of extraordinary scale, Wagner's four-part epic *Der Ring des Nibelungen* was written over the course of twenty-six years, between 1848 and 1874, at the height of romanticism. German and Nordic mythology and folklore, incorporated into the libretto of the music-drama, formed the basis of the *Ring*. The first part, *Das Rheingold* (The Rhinegold), is based on the Old Norse *Edda*, *Die Walküre* (The Valkyrie) is largely based on the *Völsunga saga, Siegfried* contains elements from the *Eddur* and *Thidrekssaga*, and the final *Götterdämmerung* (Twilight of the Gods) draws on the 12th-century German poem, the *Nibelungenlied* (The Song of the Nibelungs) - the primary inspiration of the cycle. It follows the struggles of gods, heroes, and several mythical creatures over the eponymous magic ring that grants domination over the entire world, with the drama spanning over three generations of protagonists, until the final "cataclysm" at the end of *Götterdämmerung*. The music drama was to fuse all musical, poetic and dramatic elements into the ultimate emotional synthesis, the super-unified whole of the *Gesamtkunstwerk,* where melody, poetic text, staging, and lighting all partake in telling the story. The first complete performance of the *Ring* took place in 1876 at the Bayreuth Festspielhaus (Festival), a music drama theatre designed and constructed by Wagner *ad hoc* to stage the *Ring* (Fig. 2.7). The auditorium had a single ascending bank of seats, and no boxes or circles, in order to provide all spectators with an equally good view of the stage. The theatre was also built to harbor a pit, creating a "sunken orchestra" that would conceal the musicians from the audience and enable the "merging" of various elements in the performance into a single "whole". Both egalitarian and artistically revolutionary, it revived the amphitheatres of ancient Greece, and the "artificial darkness" that Wagner "invented" for his music drama productions would later serve as a model for cinema itself (Elcott, 2016: 51).

As Wagner believed that the "mimetic" relation between aria, gesture and orchestra had to be eradicated, he "promoted" the orchestra to be on par with the singer (eventually becoming superior, in fact) through the development of a new compositional style based on the *leitmotif* (leading motif), i.e., a short, constantly recurring musical phrase associated with a particular

person, place or idea. It is defined as the "smallest structural unit in possession of thematic identity", whether melodic, harmonic, rhythmic, or all three (Albrecht and Wöllner, 2015: 145). In Wagner's music dramas, specific leitmotifs would be played as soon as a figure or object appeared on stage, enabling the spectator to know and anticipate certain events before the characters on stage. Its use was fairly widespread among composers of Romantic opera such as Carl Maria von Weber, and was closely linked to the notion of musical *idée fixe* coined by Hector Berlioz. However, Wagner adopted its use to such a scale that the *Ring* was narrated and understood almost entirely through "sonic images". Wagner interweaved hundreds of leitmotifs, many recurring through the entire cycle, to guide the listener through 15 hours of music drama. When one hears the famous "Valkyrie motif", for example, either a Valkyrie is already on the stage, about to the enter the stage, or mentioned in the *libretto* of the singers, and immediately the image of the Valkyrie is triggered in the mind of the spectator.

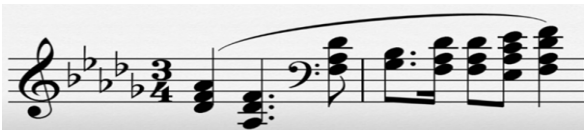
Wagner's leitmotifs were not fixed musical patterns, however, changing in relation to the plot, moving Siegfried's "Sword motif" a half-step down if there was impending peril, and up to a higher *tessitura* to denote a tender moment (Albrecht and Wöllner, 2015: 146). The characters in the *Ring* were represented by the instruments' various acoustic features, immersing the listener in a "bodily" experience of the story, a method adopted by Prokofiev in his children's symphony *Petia i volk* (1936). Known as "tone painting" (Krausshouse, 2011), Wagner would use the timbre of heavy, loud brass to mimic the weighty steps of Fafner and Fasolt, expressing their "immense strength" in the "Giant motif". Notably, the German composer went so far as to invent the "Wagner Tuba" especially for the "Valhalla motif", the home of the Gods. Even "fire" was narrated with its own leitmotif, merging in *Götterdämmerung* with the Valhalla motif to indicate the burning of the dwelling of the Gods. It is precisely in the textured, rich integration of hundreds of character and object leitmotifs into a story told musically that Wagner's genius as a dramatist is truly manifest.



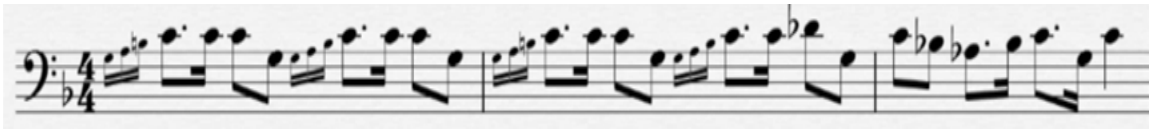
2.6

2.7

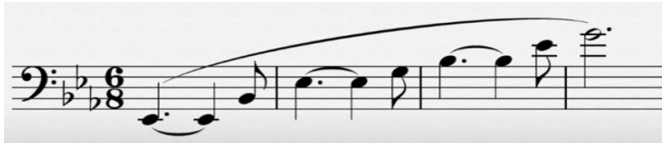
Fig. 2.6 – 2.7 | Original production of Wagner's *Ring* (Bayreuth, 1876). Bayreuth festival backstage view showing Wagner's staging techniques for *Das Rheingold*, source: *Das Rheingold* (2.6); premiere cast of the *Ring*, the Rhinemaidens - Lilli Lehmann (centre) as Woglinde, source: Zeno (2.7).



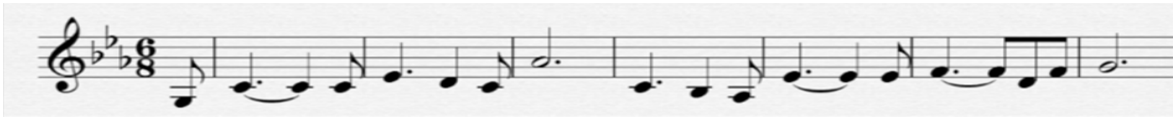
2.8



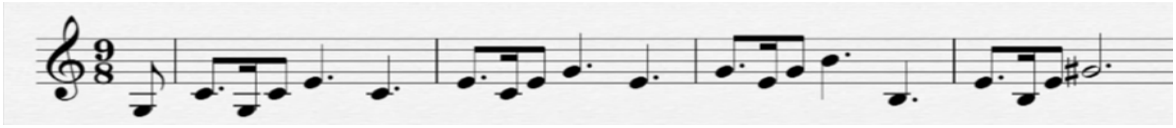
2.9



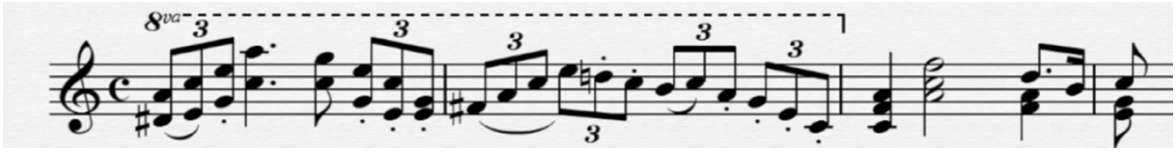
2.10



ì2.11



2.12



2.13

Fig. 2.8 – 2.13 | Wagner's musical gestures. Leitmotifs in *Der Ring des Nibelun*; Valhalla motif (2.8); Giants motif (2.9); Siegfried motif (2.10), Valkyrie motif (2.11); Genesis (Nature) motif (2.12); Heroic Love motif (2.13). Source: *Der Ring Des Nibelungen*, 2014.

As Barrlett notes, few directors that have made the transition from cinema to opera, making Eisenstein's production of *Die Walküre*, the second part of Wagner's *Ring*, a historical staging (Bartlett, 1992). From August 1939, when Ribbentrop and Molotov signed the Non-Aggression Pact, until the beginning of 1941, the Soviet government officially pursued a cultural policy designed to foster links with Germany. Following the Pact, Eisenstein was invited to stage a *Die Walküre* for the Bolshoi Theatre, premiering on November 21st, 1940. Although Wagner's music enjoyed immense popularity following the Revolution of 1917 (due to Wagner's participation in the 1849 Dresden Uprising and the circulation of his anti-capitalist brochure "Die Kunst und die Revolution" (Art and Revolution, 1849)), toward the 1920's approval for Wagner and his music waned (Bartlett, 1992: 55). Furthermore, Wagner's anti-Semitism and his status as Hitler's favorite classical composer solidified the appropriation of his artistic project by the politics of the extreme-right. Indeed, Eisenstein's involvement in the production has often been viewed as evidence for his "conformism" and "faithfulness" to the Soviet authorities. However, in the midst of the contemporary re-assessment of Eisenstein's legacy facilitated by new research and increased access to archives, Eisenstein's *Die Walküre* has been re-considered as an "anti-fascist treatment" of a subject appropriated by the Nazis (Kleiman, 1993: 39). Indeed, there are several apocryphal accounts that Nazi diplomats attending the premiere were appalled at the production (Bartlett, 1992: 56).

Eisenstein's writings on Wagner and his production plans for *Die Walküre* demonstrate the the two artists' shared interest in the link between the syncretic unity of prehistoric society, gesture, and holistic aesthetic experience. The task allotted to the director staging Wagner's work, or any music-drama, is to "catch that unifying general gesture – the thing as a whole [...] that lies under the music" and is "expressed through it" (Eisenstein, 2004: 185). In his 1940 essay "Voploshchenie mifa" (Incarnation of Myth, see: Eisenstein, 2010c), Eisenstein proposes that whereas Wagner's *Tristan und Isolde* (1865) expresses "intellectuality – thought", the *Ring* expresses "visual quality – action, deed, gesture" (Eisenstein, 2010c: 154). Eisenstein defines this as the *zrimaia osiazaemost'* (visual tangibility) of *Die Walküre*, a music that "wants to be visible, to be seen", whose climax is the "whirlwind of the 'Flight of the Valkyries'" and whose "dynamism [is] resolved scenically, vertically upward" (Eisenstein, 2010c: 166).

In his article "Pered prem'eroi 'Val'kirii'" (Before the Premiere of Valkyrie, 1940) Eisenstein writes:

> "People, music, light, landscape [...] colour and movement [...] all brought together by a single piercing emotion, a single theme and idea - this is what the film maker strives to achieve, and the producer finds the same when he becomes familiar with Wagner's works" (Eisenstein, 1940, cited in Bartlett, 1992: 60).

In Wagner's unification of the ideals of nationhood and convergence in the arts, Eisenstein found a formidable outlet for his ambitions of achieving theoretical and practical excellence in film. Indeed, it appears that the Soviet director deeply admired Wagner, praising both his socialist values and artistic genius:

> "Wagner, who saw an artist's greatest achievement as the ability to realize a people's creative will; Wagner, who asserted that a great, real, genuine work of art could not be the work of one artist alone; Wagner, who wrote that 'the tragedies of Aeschylus and Sophocles were the work of Athens as a whole'. Wagner himself, in his four part cycle *Der Ring des Nibelungen* did for the discrete, epic legends of the Germanic peoples what Homer did for antiquity with *The Odyssey* and *The Iliad*, or what Dante did for the early renaissance with *The Divine Comedy*." (Eisenstein, 2010c: 142-143).

In Wagner's work, Eisenstein also recognized the "emotional complex that connects primitive man with the whole diversity of natural phenomena" (Eisenstein, 2010c: 145). Nature's "participation in human affairs", the relic of antiquity lost by modern man, can be found again when recreating myth. As Eero Tarasti notes, Eisenstein's study of Wagner's use of myth and his notion of *Gesamtkunstwerk* can be regarded as a reconstruction of "primal mythical communication" (Tarasti, 1968: 14). Eisenstein writes:

> "The idea running through the whole spectacle – the synthetic merging of emotion, music action, light and colour – here perfects the image of "non-indifferent nature" as it appears to man's imagination in the process of creating legend and myth" (Eisenstein, 2010c: 161).

There could be no "ordinary separation of artists", the production had to "*embody the content within its actual plastic form*" (Eisenstein, 2010c: 166). Committed to honoring Wagner's innovative "production principles", for the famous "Ride of the Valkyrie" scene Eisenstein had wanted to "envelop" the auditorium in a system of loudspeakers that would "reverberate as if in flight" the iconic scene, lamenting that alas he was unable to "overcome the traditions of the opera theatre" (Eisenstein, 2010c: 167). Even before the invitation to produce *Die Walküre*, Eisenstein had a devoted interest in "myth *avant la lettre*", including the Nibelung legends. Quoting Wagner's biographer John F. Runciman, Eisenstein writes:

> "It is charged with a sense of a strange remote past – a past that never existed. No archaic chords or progressions occur, but by a series of miraculous touches the atmosphere of a far-away past is kept before us" (Eisenstein, 2010c: 161).

For the Soviet director, the project presented the opportunity of returning to the theme of the universality of myth both conceptually and in the visual form of a theatrical production (Bartlett, 1992: 53). Wagner's ash-tree was of particular interest to Eisenstein and formed the unifying basis of his interpretation of *Die Walküre*. He saw it as the ancient Yggdrasill of the Edda legends, the cosmic tree that was central to the world and the origin of everything, its branches encompassing both Heaven and Earth. After the apocalyptic destruction of the gods, the new world would be created by those who had sheltered inside the tree. Signifying the ultimate unity between mankind and nature, Eisenstein recognized that the image of the "God-tree" was basically common to all ancient mythologies, and compared Wagner's ash-tree to Dante's tree symbolism, writing:

> "[there is] something Dantesque! [sic] in this tree full of characters! N.B. And this is no accident. Wagner writes to Mathilde Wesendonck, London, 30 Aug 1855: At present I'm reading a canto of Dante every morning before I set off to work: I'm still stuck deep in Hell; its horrors accompany my realisation of the second act of [Die] Walkure [...]'" (Eisenstein, 2010c: 188).

As a key pillar of the embodiment of myth in *Die Walküre*, Eisenstein evokes the legacy of Wagner's fervent anti-capitalism. Indeed, Wagner was a radicalized participant in the European revolutions that broke out in 1848. Acquainted with revolutionary anarchists such

as Mikhail Bakunin, Wagner was committed to a vision of a world where the power of capital would be undermined, and class and other social institutions would finally disappear. In *Opera und Drama* (1851), Wagner asserts that the notion of property is not the basis of good order but rather the origin of "all the crimes in myth and history". Wagner began writing the *Ring* in the wake of the failed Dresden uprising and his eventual exile, and the major themes of love, greed and betrayal run through the work as a reaction to the lost revolution, finding an ultra-nationalist refuge in Germanic and Nordic mythology. Alberich the dwarf renounces love and plunders the Rhinegold in order to unleash the ring's power, enslaving his fellow brethren, and in the end the destruction of Valhalla and the Gods follow as the old world order, ruled by greed, collapses. Indeed, George Bernard Shaw famously argueS that Wagner's *Ring* is a socialist critique of industrial society and its abuses (Shaw, 2010). Eisenstein appreciated Wagner's "hatred for private property", but noted that although Wagner "penetrates the secret" of the myths of ancient times, the actors in his original production at the Bayreuth Festspielhaus are "far removed from the people of the myth". Eisenstein posits:

> "A bourgeoise person of the capitalist nineteenth century was always alone, always isolated, living in hostile insularity from his neighbor and his surroundings, out of harmony with the world and in irreconcilable discord with human society. That is how these people looked, dressed in their plumed helmets. They underlined their hopeless insularity from their stage surroundings with the real fur of their costumes and their weapons gleaming like household cutlery. What I wanted to achieve in my plastic treatment of this spectacle was to overcome this rupture, to have, in contrast, *the unity of the original harmony* between man and his surroundings" (Eisenstein, 2010c: 164).

The Soviet director's critique of Wagner's failure to embody his anti-capitalism through his actors and their gestures (*Gebärden*) on stage resonates with Agamben's account of the bourgeoise's gestural crisis in the 19th century. Wagner, a predecessor to cinema and the kinesthetic turn, is unable to realize Benjamin's bodily collective innervation on the stage, rendering all gesture static, unexpressed. Indeed, Eduoard Dujardin, editor of the *Revue Wagnérienne* in Paris, observes that in the scenes with Wotan, the chief of the Gods, in particular, "nothing falls to the actors, with the decorations and the gestures being in the orchestra" (Dujardin, 1887: 172). Similarly, Stéphane Mallarmé notes that the stage remains

"abstract in and of itself, impersonal", the actors "depersonalized mimes" and only the orchestra "penetrating and engulfing" the drama (Mallarmé, 1897/1976: 171). However, the vitality of Wagner's art *is* felt in the embodied expressiveness of his *music,* through which "gestural infatuation", to use Agamben's term, is mediated through melodic images. The *Ring* chronicles Wagner's desire to end the alienation of the bourgeoise through the transformative political power of embodying myth musically, transporting the listener to a time when man was not only in harmony with his own gestures, but those of nature and the Gods in a unifying whole, albeit falling short to do so with his "static" actors. The failed revolution of 1848 is revived on the stage, with the masses gathered now in Wagnerian darkness rather than in city squares. As we will see in the next section, however, the Wagnerian leitmotif and *Gesamtkunstwerk,* and the cultural institutions and artists they influenced, are prominently antithetical to Adorno's, Horkheimer's and Eisler's critical theory.

2.3 Adorno's critique

In *Versuch über Wagner* (In Search of Wagner, 1952/2005), Adorno offers his foundational critique of Wagner's aesthetic program of the *Gesamtkunstwerk*. In Wagner he identifies a "deeply bourgeois" individual unable to escape the pitfalls of the "dialectical process of enlightenment" that defined his era. Although Adorno conceives of Wagner's anti-Semitism as part of the "artistic texture" of his work, in the 1950s and 1960s he would adopt a more positive perspective on Wagner, particularly with respect to *Tristan und Isolde*. Indeed, Adorno's criticism set a new standard in the field of Wagner studies, combining a Marxist reading of musical art with meticulous musicological analysis (Žižek in Adorno, 2005: viii). As a central thesis he argues that Wagner's music as "monument to betrayed revolution" failed to counter the "formal elements" of the very "commodification" that he fought against. In his chapter on "Gebärde" (gesture), Adorno writes:

> "Among the functions of the leitmotiv can be found, alongside the aesthetic one, a commodity-function, rather like that of an advertisement: anticipating the universal practice of mass culture later on, the music is designed to be remembered, it is intended for the forgetful (Adorno, 2005: 21).

In addition to the leitmotif's primitiveness, and its implicit condescension, Adorno criticizes the "amalgamation of materials" and "primacy of effect" in the aesthetics of the *Gesamtkunstwerk* (Hansen, 2011: 209), asserting that:

> "The idea of uniting all the arts is itself dilettantish and, in the absence of the supreme effort entailed in subjecting them all to his overwhelming genius for expression, it would have remained at the level of dilettantism. There is something dubious about his relations with the arts; insane though it sounds, there is something inartistic about it" (Adorno, 2005: 19).

A notion Adorno developed together with Benjamin in the 1930s, he considers the *Gesamtkunstwerk* the embodiment of *phantasmagoria,* where the "outward appearance" of human labour involved in the collaborative work of art (such as the concealed musicians in the sunken pit) is eliminated, disguised and "dangled magically" before the unsuspecting spectator (Deathridge, 1983: 82). Musically, Adorno asserts that Wagner creates a "vertical and horizontal process of shrinking" that gives the impression of "endless space" within the "allegory of the myth". Wagner's phantasmagoric blurring of the lines between individual and collective components, isolating the audience from the mode of production, submerges the listener in the "mirage of eternity". The ills of capitalism and technology, therefore, are replicated in what is intended to be an anti-modernist opus. Achieving an "unprecedented" division of labor in the history of music, Wagner's "technological wizardry" envelops the spectator "in the self-same alienation against which the Gesamtkunstwerk rebels" (Adorno, 2005: 110).

In *Dialektik der Aufklärung* (Dialectic of Enlightenment, 1944/2002), written in American exile between 1941 and 1944, Adorno and Horkheimer scrutinize the Hollywood "culture industry", the heirs of 19th century phantasmagoria, as a system of "secondary exploitation" that "subordinates" all aesthetics to one "ideological" aim: to "reproduce the spectator/listener as consumer" (Hansen, 2012: 207). The authors assert that the true agenda of such an institution is to provide "illusory escape" from the "alienation" underlying everyday life, where "leisure time" is used to restore strength spent on the "alienated process of labor". The capacity to draw attention away from the "cinematic apparatus", and thus the awareness that the film is a "product", drives the "apparent transparency" of the diegetic world of Hollywood films. The

gesture that was revitalized by cinematic screens at the turn of the century now serves only to "distract" the spectator from his ongoing isolation and drudgery, and the primary "mode of reception" that Hollywood cinema engenders in its audience is one of "passivity". The authors assert:

> "Amusement under late capitalism is the prolongation of work [...] What happens at work, in the factory, or in the office can only be escaped from by approximation to it in one's leisure time. All amusement suffers from this incurable malady. Pleasure hardens into boredom because, if it is to remain pleasure, it must not demand any effort and therefore moves rigorously in the worn grooves of association. No independent thinking must be expected from the audience: the product prescribes every reaction. Even the set pattern itself still seems dangerous, offering some meaning - wretched as it might be - where only meaninglessness is acceptable" (Horkheimer and Adorno, 2002: 109).

This passive subjectivity and the growing difficulty in distinguishing the difference "between movies and real life" would be further enhanced by the advent of synchronized sound and the refinement of mixing technology. The sound film, "far surpassing the theatre of the illusion", forces the spectator to "equate it directly with reality". Any "sustained thought" would only distract the spectator from the "relentless rush of facts", and an unmitigated absorption into the "images, gestures and words" of the sound film leaves the spectator unable to "supply what really makes it a world" (Horkheimer and Adorno, 2002: 126).

In *Composing for the Films* (1947/2007), Adorno together with Eisler explore how the nature of music and human hearing is "exploited" for the purposes of this "narcotization". The text was written while Adorno and Eisler were employed on several research projects funded by the Rockefeller Foundation, the "Princeton Radio Research Project" and the "Film Music Project", respectively. Although the initial aim of the collaboration was practical experimentation rather than "description and prescription" (Rosen, 1980: 162), in *Composing for Films* Adorno and Eisler explore films in both "aesthetic" and "creative" terms in one of the first critiques of the role of film music in enhancing the diegetic transparency of Hollywood films. The authors introduce the text with the following formula: "The popular messages conveyed by this industry must not be conceived as an art originally created by the masses.

Such an art no longer exists or does not yet exist" (Adorno and Eisler, 2007: xxxv). With its "adaptation […] to the immediate needs of the film industry" and "pseudo-tradition", film music cultivates its own "trivialization". The authors cite several vehicles of impoverishment in the Hollywood film score, such as the "demand for melody", "unobtrusiveness" (e.g., background music), illustration (i.e., "Birdie sings, music sings"), and clichés (e.g., continual emotional crises and traditional approaches to generating suspense). Wagner's leitmotif, adopted by Hollywood music composers as a central "structuring principle" of narrative, is evoked as a primary culprit in "maintaining" spectator passivity while simultaneously "denying music its potential" (Davison, 2004: 23). Adorno and Eisler refer to the leitmotif as the "most elementary means of elucidation, the thread by which the musically inexperienced find their way about" (Adorno and Eisler, 2007: 2). Max Steiner, in contrast, committedly and explicitly relied on leitmotifs in his Hollywood film compositions. In a 1967 interview with Myrl Schreibman, Steiner asserts that a film score must serve the "dramatic content" of a film, an approach his critics call "Mickey Mousing" (a technique used in animated cartoons and what Adorno and Eisler refer to as "illustration"). Steiner declares that:

> "Gone with the Wind [Fleming, 1939], like all my scores, is written like an opera. […] If you listen to Wagner's *Ring* you will find the same theme throughout […] On screen the music has to fit, otherwise how are you going to play? If you are in the middle of the love scene, you cut away to a barroom someplace where […] what are you going to do? Keep on with the love theme? […] You have to play the appropriate music to develop the action. If that is 'Mickey Mousing', it's alright with me […]" (Schreibman and Steiner, 2004: 46).

Adorno and Eisler, however, go so far as to describe the leitmotif as a "musical lackey" who introduces his master with "an important air" despite the fact that he is "clearly recognizable to everyone". The authors argue that although music and image must "correspond" even if "antithetically", music should not "confirm" or "duplicate" action on the screen in an "immediate" way. Adorno and Eisler note that whereas in Wagner's work the leitmotif genuinely provides the audience with essential "signposts" throughout the "gigantic dimensions" of his music dramas, the essential relation between "vast dramatic canvas" and "atomization of the music element" is utterly missing in Hollywood cinema. Furthermore,

when Wagner's tubas introduce the Valhalla motif, it is not merely to indicate the home of Wotan and the gods, but to "connote the sphere of sublimity", to connote "symbolism", which "has no place" in the Hollywood motion picture which strives to represent only "reality" (Adorno and Eisler, 2007: 3). Adorno's notion of an "alternative" music originates in the ideas of the Second Viennese School, such as those of Arnold Schönberg and Alban Berg, whereas Eisler's contribution to the volume was founded on his practical experience of composing music in Weimar films such as *Kuhle Wampe* (Dudow, 1932), collaborating with Brecht on plays such as *Die Maßnahme* (1930) and *Die Mutter* (1931), and taking part in Hollywood productions such as *Hangmen Also Die* (Lang, 1943).

In addition to the practices of Hollywood, Adorno and Eisler also draw into question the soundness of Eisenstein's analogy between cinematic and musical movement, asking "why should one and the same thing be reproduced by two different media?", and asserting that "the effect achieved by such repetition would be weaker rather than stronger" (Adorno and Eisler, 2007: 45). The authors argue that Eisenstein is not "altogether free from the formalistic type of thinking he so properly attacks", that such "parallelism" between image and sound would be merely the transfer of principles of "pseudo-identity or association" to a more "abstract level", which contrasts with Eisenstein's "rejection" of "absolute equivalence" (Adorno and Eisler, 2007: 45). The authors posit: "[…] the aesthetic divergence of the media is potentially a legitimate means of expression, not merely a regrettable deficiency that has to be concealed as well as possible" (Adorno and Eisler, 2007: 52). Although the authors agree with Eisenstein on the role of rhythm in pacing the movement in both image and sound, Eisenstein's formula that musical movement must be grasped as the "foundation" for corresponding plastic composition, Adorno and Eisler maintain, is a manner of thinking that is "both too narrow and too vague". The authors write:

> "If in the name of higher unity, picture and music were made to present this rhythm incessantly and simultaneously, the relations between the two media would be pedantically restricted, and the result would be unbearable monotony. Movement can also mean a higher aesthetic quality of the motion picture; and it is this quality that Eisenstein obviously has in mind" (Adorno and Eisler, 2007: 46).

Nevertheless, drawing also on Benjamin's and Brecht's views, Adorno and Eisler agree with Eisenstein's ethos that montage should be the basis for the practice of filmmaking, allowing for both independence of the individual arts and unification of constituent parts where needed. As Nicholas Cook observes, as long as "Romantic aesthetics persist", the "alliance" of one medium with another is "bound to attract suspicion, like the old formula that a truly genius work of writing or music cannot be 'contaminated'". As Rainer Maria Rilke declares, "I hate to believe that there could be any room left over for another art" (Rilke, 1969, cited in Cook, 1998: 128).

* * *

In the next chapter, the link between gesture and multimodality will be examined morely closely through Eisenstein's writings on organic unity, regression, and pathos, and the unifying gestures of "rotating" and "rising" in *Ivan Chast' II* and in the films that followed Eisenstein.

3. The audiovisual gesture

In *Neravnodushnaia priroda* (Nonindifferent Nature, 1940/1988), Eisenstein writes:

> "Genuine composition is immutably, deeply human: whether this be the bouncing rhythm of [...] cheerful episodes, the monotonously drawn-out quality of the editing of a sad scene, or the joyful glittering in the lighting of a shot" (Eisenstein, 1988: 23).

In this chapter, I will discuss: 1) d'Udine's ideas on gesture and synesthesia in art and Eisenstein's ideas on organic unity, regression and pathos; 2) the manifestation of these themes in the "oprichnina scene" of *Ivan Chast' II* (1958); and 3) an analysis of multimodal gestures outside the cinema of Eisenstein, notably in *The Red Shoes* (The Archers, 1948), *Moi laskovyi i nezhnyi zver'* (The Hunting Accident, Loteanu, 1978) and *The Shawshank Redemption* (Durabont, 1994).

3.1 Rhythm, organic unity, and *pathos*

As mentioned in Chapter 1, d'Udine's *L'art et le geste* (1910) is a somewhat forgotten source of inspiration for Eisenstein's ideas about synaesthesia in art (Schulzki, 2019). D'Udine conceived of a central thesis that can be found on the cover of his book: "all artistic creators are specialized mimes". Expanding on this formula, d'Udine writes:

> "[...] each emotion, of any order, corresponds to only one attitude or bodily movement, and that it is through the intermediary of this movement that the extremely complex synesthetic translation which accompanies all artistic creation takes place" (d'Udine, 1910: 63 (my translation)).

According to d'Udine, the origin of all "artistic impulse" lies in rhythm and gesture. Nature leaves its "trace" in the body, which in turn "expresses the impulse" in the form of vibration ("Vivre c'est vibrer"; d'Udine, 1910: 3). All the arts – music, sculpture, poetry, painting and architecture – are united by a common "motor basis". Touch, d'Udine asserts, is the primary faculty of perception and is the frame of reference for the other human senses of audition, smell, sight and taste, an idea that resonates with the theory of "Gibsonian affordance" that would dominate psychology a half-century later. Similarly, it echoes Freeberg and Gallese's

(2007) notion of the "implied gesture" of the trace of brushwork in a painting. Calling it a "curious phenomenon" among writings on aesthetics, Eisenstein praises d'Udine's text and cites it frequently in his essay "Opredeliaushchii zhest" (The Underlying Gesure, 1939-1940; see: Eisenstein, 2004). Reflecting on the commensurability of sound and color, Eisenstein writes:

> "Their connection is in the emotional stroke of movement – in the gesture of the artist, which is the basis of all artistic manifestations. [...] in place of the famous dictum of Hans Bülow, "In the beginning there was Rhythm", we will use a more precise and together a more general formula: in the beginning there was Gesture" [Jean d'Udine] 1910" (Eisenstein, 2004: 183 (my translation)).

Although d'Udine does not refer to cinema (as Schulzki notes, it was not yet considered an art form), Eisenstein asserts that:

> "The book was published long before the full triumph of cinema [...] but it is in cinema, as perhaps nowhere to such an extent, that this principle [gesture] and the principle of synaesthesia are confirmed. [...] this factor of gesture as the primary defining impulse to create an image [...] it is the gesture that is the key to create a synesthetic unity between image and sound" (Eisenstein, 2004: 175 (my translation)).

Going further, Eisenstein calls it the "primary embryo of expressive form, from which emotion is poured out which [...] grows into the fullness of the image, plastic or sound, color or musical, or uniting all of them in the audiovisual image" (Eisenstein, 2004: 175 (my translation)). Eisenstein elaborates his thesis of "organic unity", where music, language, and color are "commensurate" elements in the pre-differentiated "language" of "rational" and "pre-rational" layers of consciousness. Eisenstein writes:

> "When there is still no strict differentiation between different types of perception. When one is still able to see [...] sound and hear [...] colors [...] "screaming" colors or "transparent" music [...] "deep" tones in painting or "variegation" in the orchestra" (Eisenstein, 2004: 170 (my translation)).

In his text, d'Udine dedicates particular attention to the sister arts of music and dancing. All music is "synesthetic in character", d'Udine explains, especially "emotional music", which he considers to be "double synesthesia". The musical expression of any emotion derives from the "physiological rhythms" experienced when feeling the emotion. Ultimately, it is in the emotional reaction of the audience that the true test of art lies, d'Udine notes, and in this, the "Wagnerian art" was so successful that for twenty or so years "all French music became Wagnerian" (d'Udine, 1910: 263). D'Udine posits that it is by virtue of a "reversible phenomenon" that music became "the mother of dance". Although music appears to be the "natural" source of dance, historically speaking, it is dance that "engendered" music when the first "conscious representatives" of our species discovered the rhythms of their movement. For d'Udine, dance is "silent rhythm", a music that we can "see" (d'Udine, 1910: 11). Evoking the theme of lost gesture in the 19th century, d'Udine observes that if man were to be "freer" and "obey his instinct", dancing to the classical music he hears, the musical ideologues would call it "sacrilege" (d'Udine, 1910: 56). In "Opredeliaushchii zhest", Eisenstein cites Wagner's own recognition of the primacy of dance, "the composer who said that 'the musical and poetic arts become understandable only through the art of dancing'" (Eisenstein, 2004: 184). It is not only the rhythm of music that guides the moving body, d'Udine notes, but also the melody that affects our "muscular reflexes". For example, when we say that a certain waltz is more "danceable" than another, we are alluding to the notion that one melody is more evocative of "whirling" or "pirouetting" than another, even though the choreography of the waltz remains the same (d'Udine, 2010: 58). Notably, the gesture of "whirling", or "rotation" is one that permeates through Eisenstein's ideas. He evokes this motif when comparing Dante's circles of Hell and Paradise to the "concentric circles" of the branches of Wagner's ash-tree (Eisenstein, 2004: 188), and, as we will see, in his writings on organic unity and regression, and the *mise en scène* of *Ivan*.

A central idea of Eisenstein's opus *Metod*, a project Eisenstein began working on after his return from Mexico in 1932, was that every artistic form could gain emotional force by reactivating "lower", or "earlier", strata of consciousness and culture (Somaini, 2016: 31). In this text, Eisenstein seeks to illustrate how artistic forms need to attain a "dialectic polarity" between "regression" and "progress" in order to engage the spectator emotionally in a powerful and profound way. The effect of a work of art, Eisenstein claims, depends on two simultaneous processes: one a determined progressive "ascent" towards ideas at the "highest

peaks of consciousness", and another a "penetration" into the "deepest layer of emotional thinking" (Somaini, 2016: 32). It is what Eisenstein refers to as the *Gründproblem* (fundamental problem) of art - the paradoxical coexistence of the logical and the sensuous, the cognitive and the emotional, the conscious and the unconscious in the work of art. The "laws" of sensual thinking that first emerged in earlier stages of human development still exist on an unconscious level in modern man, and the *method* of art, according to Eisenstein, lies precisely in tapping into "pre-rational thought". One of the principal mechanisms of descending the spectator into earlier forms of thinking in Eisenstein's "method" is through what he terms the "rhythmic drum" (*drob' barabana*). Echoing d'Udine, Eisenstein regards rhythm as the essence of all organic beings, a tool for immersing the observer into the most "recipient" state possible to "receive" the work of art. In the course of his exploration of the "universality" of rhythm, Eisenstein noted the reoccurrence of certain "rhythmic" patterns inherent in nature in the praxis of certain cultural traditions, such as the whirlwind of the eastern dervishes, Mexican dancers who "keep the same rhythm from morning to night", and the rhythmic cries of crowds honoring Our Lady of Lourdes. It is particularly apparent in the operations of "magic" and "ritual," such as the "shamanic" trances in Siberia that "hypnotize" the audience with the incessant repetition of simple rhythms. When this "rhythmic drum" is activated, a "total rejection" of rational thinking occurs, and the mind "regresses" to a state where the "everything with everything connection [...] characteristic of primitive thinking" is activated. Cinema, notes Eisenstein, through its "organic expressive means" is particularly amenable to enable this "predisposition" to "sensory thinking" in the viewer. Eisenstein writes: "Hence the 'attractive' strength of film. Hence the colossal impact of this least tangible and corporal of the arts (I do not detach film from music)" (Eisenstein, 2002: 192 (my translation)).

In *Neravnodushnaia priroda* (1940/1988/2004), Eisenstein explores the historical manifestations of his notion of *pathos* and *pathos composition*. The path to this aesthetic principle, Eisenstein explains, lies in the transition from external, object-focused perception toward an "internal tonality" characterized by a primacy of feeling. Whereas for Plato *pathos* (i.e., affect, or passion) that is "unrestricted" leads to a path of "destruction" and "contagion", for Eisenstein art *must* turn into "pure pathos" (Iampolski, 2016: 370). As Somaini notes, Eisenstein imagines "an ecstatic 'flow'" that links the artist, the work of art and the spectator in a stream of "dialectic, ecstatic" energy in a "continuous process of becoming" (Somaini, 2016: 70). For Eisenstein, *pathos* correlates closely to the notion of "ecstasy", or "ex-stasis", a

feeling of being "outside of oneself" where a certain "transition" occurs – from sound to image, from words to color, from the internal to the external (Eisenstein, 1988: 154). It is through *pathos* that the gap between representation and that which is represented must be filled, by "using the structure of human emotions as the source, since it unmistakably appeals to one's emotions, and unmistakably evokes the complex of emotions that engendered it" (Eisenstein, 1988: 5). In cinema, a "production will become organic and achieve unity – in the domain of pathos as we understand it – when the theme of the work of art, its content and ideas, became an inseparable, organic whole with the thoughts, feelings, being and existence of the artist" (Eisenstein, 1988: 36). Through "emotional isomorphism" with the artist's creative process, the artist's emotional state re-emerges in the embodied experience of the beholder, with art itself emerging "at a site of emotional communion between artist and spectator" (Bordwell, 2005: 190). Through "dialectical unity" of contradictory elements and the law of transformation from "quantity to quality", "leaps" into new states occur – such as the spectator who "jumps out of his seat" or "he who must applaud, cry out" (Eisenstein, 1988: 27). The image of the whirling Dervish, Eisenstein explains, is a "crude" form of ecstasy constructed purely on the basis of the "pulsating principles of rhythmic gymnastics" - a "prototype" of imitative behavior, such as the figure that behaves "ecstatically on the screen [...] seized by pathos" (Eisenstein, 1988: 55). It is necessary, Eisenstein explains, to "go beyond the limits of the human being, spreading [...] to the surroundings of the character" (Eisenstein, 1988: 29). It is necessary to transition "from the dramatic to the metaphoric", like "water that transforms into vapor" and "jumps" to a new "sensory quality" (Eisenstein, 1988: 50). In cinema, this transition occurs when there is "a leap from [...] acting (theatrical cinema) to [...] 'pure' cinematography, of [...] unmatched means and possibilities" (Eisenstein, 1988: 23). In comparison to art, the euphoria of religious experience pales as it receives "divine colouring post-factum" (Eisenstein, 1988: 177). Only when the artist himself is "overcome by the pathos of theme" can organic unity and the ecstasy of artistic creation occur. This aesthetic exultation, however, must be grounded in the "basis of the mutual play of human emotions [...] and human experience", what Eisenstein calls the "human approach" (Eisenstein, 1988: 4). For Eisenstein, the prototype of aesthetic composition will always remain the "actions of man" and "the structure of human actions" (Eisenstein, 1988: 8). As Somaini notes, Eisenstein's notion of *pathos* shares a notable correspondence with Aby Warburg's notion of *Pathosformel* (Somaini, 2016: 61). Warburg's concept refers to the existence of a few "formulas" of bodily

gesture inherent to the expression of a set of emotions (e.g., suffering, mourning, Dionysiac pathos, fury) that are deeply ingrained in the cultural memory of their migrations throughout the "history of images" (Somaini, 2016: 64). Eisenstein identifies similar cultural patterns in El Greco's *La Resurrección de Cristo* (Resurrection, 1600), Pablo Picasso's *Toro moribundo* (The Dying Bull, 1934), Giovanni Battista Piranesi's *Le Carceri d'Invenzione* (Prisons, 1750), the classic figure of the "adulteress" in Leo Tolstoy's *Anna Karenina* (1877), and in his own films, such as the "Odessa steps" sequence in *Potemkin* (1925) and the "milk separator" scene in *General'naia Liniia* (The General Line, 1929). Most notably, Eisenstein's later theories on synthesis in film find their most complete expression in *Ivan Chast' II* (1958).

3.2 The "oprichnina" scene in *Ivan Chast' II* (1958)

In many ways, Eisenstein's last film *Ivan Chast' II* is the apotheosis of his ideas on organic unity, pathos and regression, and a tribute to the Wagnerian *Gesamtkunstwerk* through the medium of film (Peatman, 1976; Thompson, 1982; Khitrova, 2011; Tsivian, 2019; Neuberger, 2003, 2019). Eisenstein asked Prokofiev to compose a Wagnerian score incorporating numerous motifs in order to musically orientate the spectator in what was meant to be a three-part saga. Indeed, Prokofiev's score for *Ivan* features a far "richer orchestral pallette" than *Nevskii*, layered with Russian liturgical music and complex melodic and harmonic arrangements. As Khitrova (2011) notes, most of the actors' movements in *Ivan* are easier described in terms of choreography, and it is a testament to the kinetic qualities of Prokofiev's music and Eisenstein's figurative *mise en scène* that in 1973 *Ivan* was adapted into a ballet by composer Mikhail Chulaki and choreographer Yuri Grigorovich. It is in the climactic finale of *Ivan Chast' II*, at the nocturnal feast and dance of the *oprichnina* (Ivan the Terrible's secret police), that gestural expressiveness "erupts into dance" (Khitrova, 2011: 66).

It was well known at the time that Stalin, like Ivan, enjoyed banquets of entertaining Song and Dance, and used to invite the Red Army Ensemble of Song and Dance to Kremlin receptions. According to the correspondence of Dovzhenko, Mikhail Romm and Vsevolod Vishnevsky, it was the "various sorts of hints" and "quaint parallels with our time" in the "oprichnina dance" that caused Stalin's ban of *Ivan Chast' II* and eventually the destruction of *Chast' III* (Khitrova, 2011: 65). As Khitrova notes, it is possible that Eisenstein had choreographed "hidden messages" that the pre-censored script was unable to express directly. Historically, the *oprichniki* order was not merely the Tsar's secret service, but a "quasi-

monastic fraternity" with its own "secret rituals" and "signs". Dance "bonds those in the know", Khitrova observes, and becomes a form of ritualistic participation. Using Agfa film stock captured from the retreating *Wehrmacht*, Eisenstein unified the explosion of color with that of "ecstatic" dance. The scene was staged to Prokofiev's prerecorded music, and using axial editing the dance was cut to the rhythm of the score in order to "fragment" the dancers' movements and gradate the "tonal prominence" of the color themes (Bordwell, 2005: 249). The first part of the sequence, called the "Chaotic Dance", is full of "flying" mid-air jumps and spirited, traditional folk "barynia" with frequent stomping and squatting (Fig. 3.1 – 3.2). In contrast, the second part, called the "Orderly Dance" (Fig. 3.3 – 3.4), takes a more sinister tone and prominently features the "whirlwind" gesture of spinning. Eisenstein uses the English word "whirlwind" to describe the "kinetic equivalent" of the "snowstorm" motif that Prokofiev included in the musical score and echoes an earlier raid by the *oprichniki*. The incessant spinning and circling performed by the dancers also serves to express the "chromo-audio-visual motif" of "burning". This is evident when Basmanov, Ivan's protégé, sings the second refrain of the brutal "Song of the Oprichniki" (written by Vladimir Lugovsky), with the other guards gathered around him:

> As the drunken guests were leaving for home,
> Yes, for home,
> They burned the boyars' houses behind them.
> Yes, they burned them.

The *oprichniks* then join in drunken, mocking "revelry" and dance to Prokofiev's "rollicking" and "bumptious" score, as the Chorus sings:

> Ho! Ho!
> Chop! Chop!
> Let those axes chop them sharply.
> Oh!
> Burn! Burn! Burn!
> Burn! Burn! Burn!
> Burn! Burn! Burn!
> Burn! Burn! Burn!
> Burn! Burn! Burn!
> Burn! Burn! Burn!

It is precisely in the multisensory choreography of this scene that Eisenstein's ideas on the rhythmic drum and regression manifest in mature form. In his essay "Magiia Iskusstva" (The

Magic of Art, 1947), which Luria conserved in his archives for fifty years, Eisenstein writes that the spectator:

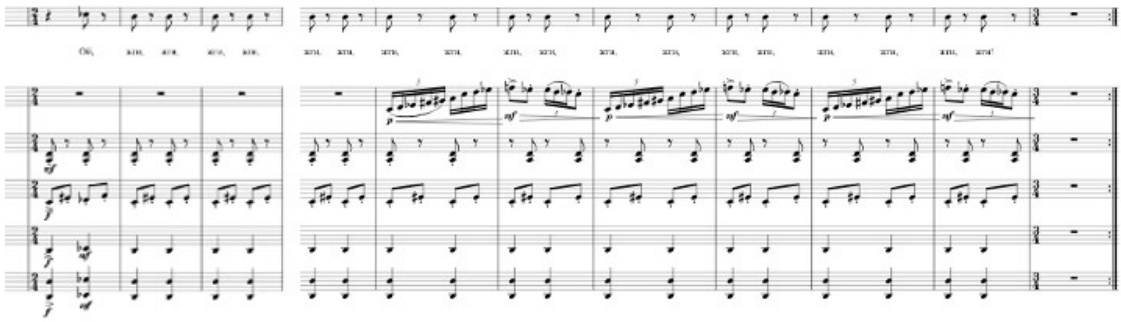
> "[…] will lose all distinction between the subjective and the objective […] where colors will sing to him and where sounds seem to be take form (synesthesia), where the inspirational word will make him react as if the fact indicated by the word had have transpired (hypnotic behavior) […] what one might call a "rhythmic drum" (Eisenstein, 2002: 185).

The exaggerated, rhythmic repetition of the word "*Zhgi!*" (Burn!) by the chorus (no less than eighteen times) feels like mystic, shamanic chanting, and the incessant spinning of the "dancing henchmen" is reminiscent of Eisenstein's aforementioned "whirling dervishes". The "whirling" of the guards causes the fabric of the costumes to "flap ecstatically" as they "glow yellow and red", a "kinetic" expression of the fire motif (Gallez, 1978: 27). This "synesthetic fusion" of what Kristin Thompson (1982) refers to as the "power gesture" in the oprichnina scene guides the viewer, possessed by the "magician-creator", into a primordial "trance". A synesthetic, Dionysian feast staged in medieval Russia, it is reminiscent of what Nietzsche, in his discussion of Wagner in *Die Geburt der Tragödie aus dem Geiste der Musik* (The Birth of Tragedy from the Spirit of Music, 1872/2003) refers to as the Dionysian *Tanzgebarde* (gesture of dance). The *Tantzgebarde* introduces corporeality into the Apollonian visual, "mobilizing within visuality the antirepresentational potential of music" as it unleashes its "rhythmic destruction" (Puchner, 1999: 29). The rhythm of "turning bodies" and "ecstatic revolving" simultaneously evoke what Marie Rebecchi, Elena Vogman and Till Gathmann (2017) refer to as the "anthropological quality of organizing experience" (Rebecchi, Vogman and Gathmann, 2017: 9) and Nietzsche's notion of the "primal pleasure" of the Dionysian (Nietzsche, 2003: loc. 179-180). The pathos-infused scene also conveys Daniel Stern's (2010) notion of "vitality effects", or the "manifestations of being alive" that are conveyed through movement. For Stern, the time-based arts of music, dance, theater and cinema "move us" by the "expressions of vitality that resonate in us" (Stern, 2010: 3).

"Correspondences" between art forms, Stern writes, "are necessarily created because of the meta-modal nature of vitality forms that assure a common ability to render similar, but not

3.1 3.2

3.3 3.4



3.5

Fig. 3.1 – 3.5 | "Dance of the Oprichniki" in *Ivan Groznyi Chast' II* (Eisenstein, 1958). "Chaotic Dance" (3.1 – 3.2); "Orderly Dance" (3.3 – 3.4); Prokofiev's Op. 116 musical score for "Orderly Dance" (written in 2/4 time; lyrics: "Oi, zhgi zhgi, zhgi! ..." (Oh, burn, burn ...!)), source: Musescore; *Ivan the Terrible*, 2016 (3.5).

identical, experiences", and the trick lies "in pairing the similar with the 'not exactly the same'" (Stern, 2010: 78). In the next section, instances of audiovisual fusion in film will be considered more broadly.

3.3 Beyond Eisenstein: multimodal rotating and rising

Michel Chion's (1994) term "synchresis" denotes the "added value" of the "spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time" (Chion, 1994: 63). According to Chion, our belief that in *Persona* (Bergman, 1966) the "shots of the hands in the prologue [...] are indeed the sounds of the hammer pounding nails into them" is the result of synchresis, a forging between "synchronism" and "synthesis" (Chion, 1994: 63). When there is simultaneity between sound and image, an "unintentional" fusion occurs between them that is difficult to unravel. Although Chion notes that synchresis is partially "Pavlovian", it is not always automatic – if one plays a "stream" of audiovisual media, certain combinations will "come together", whereas others will not. Synchresis relates closely to the notion of the "acousmatic", a term coined by the novelist Jerome Peignot and conceptualized by composer Pierre Schaeffer, which denotes the "sounds one hears without seeing their originating cause" (Chion, 1994: 72). This is, of course, the nature of all non-diegetic sound in film. However, the nature of combined media through semiotic (e.g., spatial) modes varies considerably (Bateman and Schmidt, 2013: 12). For Jerrold Levinson (1990/2011), "most (though not all) artistic entities described as multi- or mixed-media phenomena would be [...] juxtapositional hybrids", where individual media "form a whole by summation and not by merging or dissolution of individual boundaries" (Levinson, 2011: 31). After juxtapositional hybrids, there are "transformational" hybrids, when a particular art form is modified in the direction of another through the addition of specific features of that form, such as in the case of kinetic sculpture (Bateman and Schmidt, 2013: 91). Finally, there are "synthetic" hybrids, in which:

> "the objects or products of two (or more) arts are brought together in such a way that the individual components to some extent lose their original identities and are present in the hybrid in a form significantly different from that assumed in the pure state" (Levinson, 2011: 31).

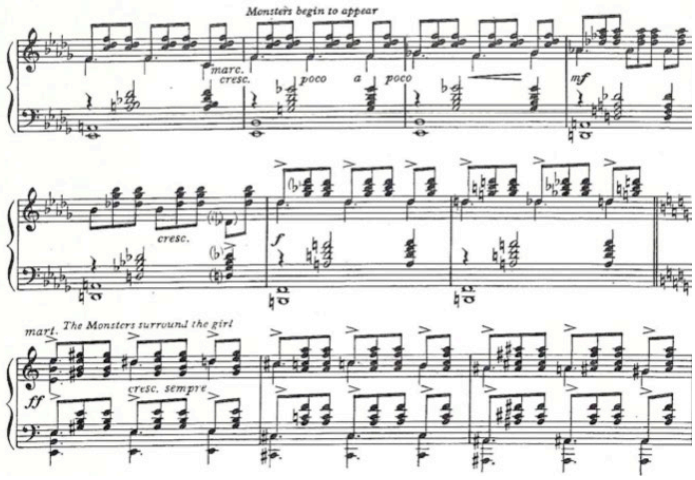
Wagnerian opera, according to Levinson, is a synthetic, fused hybrid. In light of this categorization, three scenes from world cinema are of interest here for their syncretic, synthetic hybridization of "turning" and "rising" gestures.

The first film of interest is the *The Red Shoes* (The Archers, 1948), co-written, produced and directed by Michael Powell and Emeric Pressburger. The ultimate ballet film and a "benchmark of modernity", as Ian Christie observes, "*The Red Shoes* was above all a work of creative collaboration between artists and technicians in many fields, on a scale rarely attempted in the cinema" (Christie, 2002: 83). The production included lead dancers Moira Shearer and Leonide Massine, a former protégé of Sergei Diaghilev's, as well as the choreographer and dancer Robert Helpmann, invited from the Ballets Russes. The sets were designed by Hein Heckroth, a surrealist painter trained at the Bauhaus, filmed in Technicolor by Jack Cardiff, and edited to music composed by Martin Easdale (Powell, 2000). As Diana Diamond notes, *The Red Shoes* is "a fairytale, within a ballet, within a fiction film" (Diamond, 2016: 104). Hans Christian Andersen's *The Red Shoes* (1845), the source material for the film, tells the story of a girl who is compelled to dance continually in the bewitched red shoes she formerly desired, ultimately leading to her demise. The fairytale is retold in the film as a ballet performed by the Ballet Lermonotov, a company created especially for the film. The central themes of the story are the Diaghilev-Nijinsky "mad genius" motif symbolized by ballerina Vicky (Moira Shearer) and impresario Lermontov (Anton Walbrook), as well as the conflict between art and personal life (Diamond, 2016: 107). In the film's iconic fifteen-minute "ballet" sequence, Vicky puts on the sorcerer-shoemaker's demonic shoes and begins her possessed dancing as she is transported through various landscapes that alternate between the dreamish and the nightmarish. She is led by the shoemaker to the "dead city", where Vicky is slowly surrounded by zombie-like figures (Fig. 3.6 – 3.12). Filmed from Vicky's point of view, we see as the monsters begin to approach Vicky, with Easdale's score (Fig. 3.12) initiating its *crescendo*. Because of this POV shot, when the camera pans it appears as though Vicky is *turning* her body each time, in fear and panic, as the monsters slowly walk toward her. Furthermore, the camera pans each time in metric *rhythm* to the musical score. We then cut

3.6 3.7

3.8 3.9

3.10 3.11

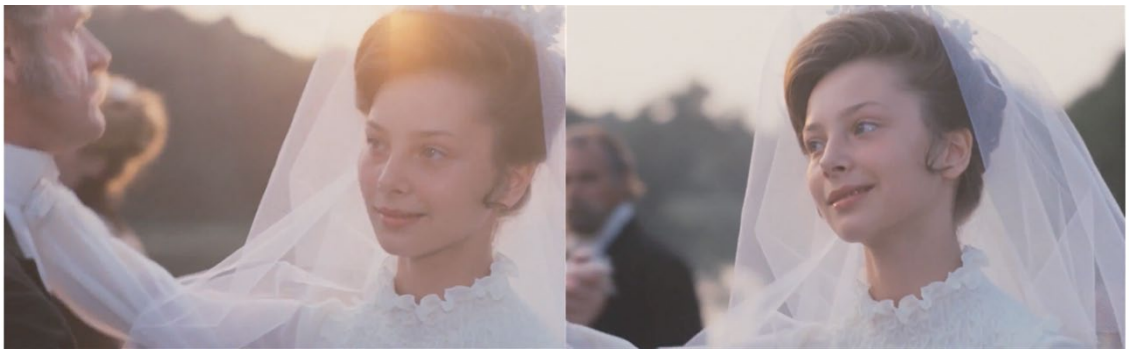


3.12

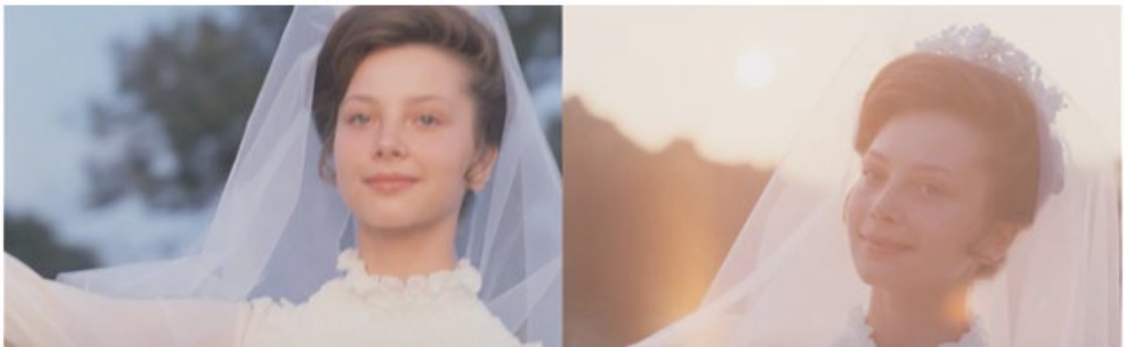
Fig. 3.6 – 3.12 | "The Dead City" in *The Red Shoes* (The Archers, 1948). Cinematography by Jack Cardiff. Camera (3.6 – 3.9) and ballerina Moira Shearer (3.10 – 3.11) synchronizing to the rhythm of Brian Easdale's film score (written in 4/4 time, source: Chappel and Co. Inc; *The Red Shoes*, 2012 [8:23 – 9:35]) (3.12).

to a wide shot of the stage, where Vicky begins to dance as the monsters torment her, with Vicky's arms now moving in rhythm to the music. Whereas in *Ivan* turning was carried out only by the repeated gestures of the *oprichniki*, an additional level is added in *The Red Shoes*. The camera/Vicky turn repeatedly as Vicky "regresses" into her own "unconscious" and "earlier layers of thought", as conveyed by the synthetic camera/body/rhythm hybrid. As Christie notes, what is essential to the cinema of Powell and Pressburger is Eisenstein's notion of synesthesia, with Powell quoting Rudyard Kipling's adage that "all art is one" throughout production (Christie, 2002: 22). Dance, music, design, art, literature, drama and cinema are united so memorably that a new art form , "choreocinema", is born (McLean, 2008: 46).

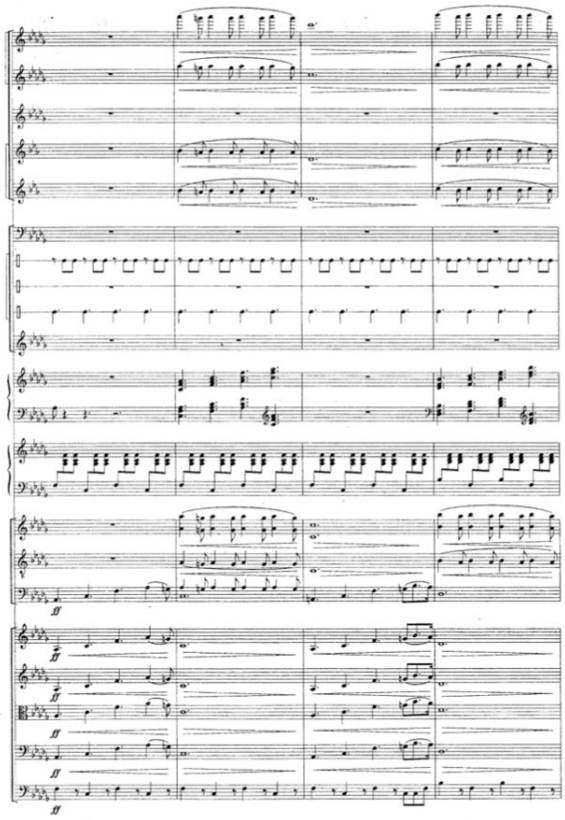
The second scene of interest is the "Wedding" scene in the Soviet film *Moi laskovyi i nezhnyi zver'* (The Hunting Accident, 1978), directed by Emil Loteanu and adapted from Anton Chekhov's 1884 novel *Drama na ohote* (The Shooting Party). Set at the end of the 19th century on an aristocratic estate in the forests of central Russia, it tells the story of Olga (Galina Beylayeva), the young daughter of a forester, and the three middle-aged men who, living in the surrounding manor, fall in love with her. In her attempt to escape poverty, she marries the nobleman Urbenin (Leonid Markov), although during her marriage she declares love to Kamyshev (Oleg Yankovsky), the court investigator. During the "Wedding" scene, the newlywed couple dance on a boardwalk overlooking a river to Eugen Doga's "Waltz". Composed especially for the film by Doga, it is a protagonist in its own right. It became famous worldwide after the film's premiere at the 1978 Cannes Film Festival, with Ronald Regan calling it the "waltz of the century" (Vechernaia Moskva, 2018). Doga's dramatic, neo-Romantic score, characterized by an ephemeral uneasiness, sets a foreboding tone for the rest of the film. As Olga and Urbenin begin to dance, the camera frequently lingers on the three onlooking gentleman, one of which is Count Karneev (Leonid Markov), who is also in love with Olga. Olga, anguished at her position, begins to cry. As the waltz approaches its climax, Olga begins to have flashbacks of her secret rendez-vous with Kamyshev, in which Olga is twirled around in his arms. From Olga's fantasizing, we cut back to the wedding, but this time Olga and Urbenin appear to be "hovering" as they are filmed in a "double dolly" shot. Urbenin's waltzing figure gradually leaves the shot frame, leaving only Olga who, "twirled" by the camera amid the dancing couples, appears to "float" away from the other characters, and sedated by the twirling and her thoughts of Kamyshev, she slips away into a trance.



3.13 3.14



3. 15 3.16



3.17

Fig. 3.13 – 3.17 |The "Wedding" waltz in *Moi laskovyi i nezhnyi zver'* (Loteanu, 1978). Cinematography by Anatoli Petritsk. Waltz between Olga (Galina Beylayeva) and the camera (3.13 – 3.16); first few bars of the "Waltz" (3.17), composed for the film by Eugen Doga (written in compound 12/8 time), source: eugendoga.musicaneo.com; Doga, 2015.

The immobility of Olga's body in contrast to the whirling of the scene, only her head gently swaying to the waltz, paints a portrait of quiet renunciation to fate – what moves her forward is not her own acting body but the now unstoppable momentum of events. Notably, there is an impressive effect created by the double dolly technique and the 12/8 time of the music. Each quarter note beat consists of a group of three eighth notes, combining the steadiness and stability of 4/4 time with the "flowing" effect of 3/4 "waltz time". Going beyond the turning bodies of *Ivan* and the turning camera of *The Red Shoes*, in *Moi laskovyi i nezhnyi zver'*, camera, actor and melody "turn" together. The effect is one of an impressive hybrid of synthetic fusion between the turning flow of the dolly and the twirling flow of the waltz – with one heightening the effect of the other. As we will see, the notion of "musical rotation" will be investigated in Chapter 6.

Finally, the "Mozart opera" scene in *The Shawshank Redemption* (1994), written and directed by Frank Darabont, will be considered. Adapted from Stephen King novella *Rita Hayworth and Shawshank Redemption* (1982), it tells the story of Andy Dufresne (Tim Robbins), a banker who is sentenced to two consecutive life terms in prison in the tough Shawshank State Penitentiary for the murders of his wife and her lover, which he did not commit. In the cruel environment of Shawshank, he befriends Ellis "Red" Redding, the narrator of the story. Despite his brutal and unforgiving surroundings, Andy fights to hold on to his personal sense of integrity and "hope", which serves as a primary theme of the story. After several vicious attacks, Andy is eventually reassigned by the warden (Bub Gunton) to the prison library, a pretext for Andy to manage the prison's finances. Andy starts to write weekly letters to the state government asking for funds to improve the run-down library, later receiving a donation that includes a recording of Mozart's "The Marriage of Figaro". Despite certain punishment, Andy walks into the Warden's office and plays an excerpt of the opera over the public address system. In this memorable scene, the inmates standing outside stare upward at the speaker system, the camera revealing the object of their fixation in an upward tracking shot. What is significant is that the camera rises in unity with the Duettino "Sull'aria".

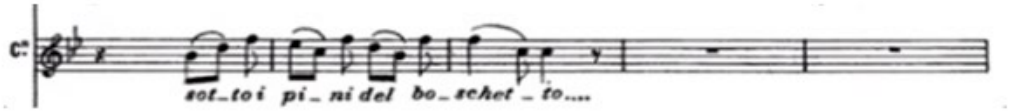
In the voiceover, Red narrates:

> "To this day, I have no idea what those two Italian ladies were singing about. Truth is, I don't want to know. Some things are left best unsaid. I would like to think they

3.18 3.19

3.20 3.21

3.22 3.23



3.24

Fig. 3.18 – 3.24 | **The "Mozart opera" in *Shawshank Redemption* (Darabont, 1994)**. Cinematography by Roger Deakins. Andy (Tim Robbins) plays the "Mozart opera"; upward tracking shot as the inmates gaze up at the speaker system (3.18 – 3.24); "Sull'aria" from *Le Nozze di Figaro* (3.24) sung by Edith Mathis and Gundula Janowitz, source: Score's Animation, 2016.

> were singing about something was so beautiful it can't be expressed in words and makes your heart ache because of it. I tell you *those voices soared higher and farther* than anybody in a Gray place dares to dream. It was like some beautiful bird flapped into our drab little cage and made these walls dissolve away. For the briefest moment, every last man in Shawshank felt free" (my emphasis, [1:09:22-1:10:04]).

The scene encapsulates a moment of *Gesamtkunstwerk* – all aesthetic elements are united in the pathos of a single feeling, a single *embodied metaphor*. The transcending effect of music, mirrored by the rising trajectory of the camera, fuse together as a liberating sense of hope is expressed through the multimodal motif of "soaring" (as Red observes). Like in *Moi laskovyi i nezhnyi zver'*, the semantic modes of the camera and melody merge into a synthetic gliding gesture. As we will see in the next chapter, the gestures of rising and spinning will be discussed in relation to the aesthetic model of this thesis – the Soviet film *Letiat Zhuravli* (1957) by Mikhail Kalatozov.

* * *

In light of the theoretical basis explored in previous chapters, Chapters 5 to 7 will focus on a "neurohumanistic" operationalization of Eisenstein's aesthetics of the "underlying gesture". Specifically, the materials, methods and results of an interdisciplinary investigation of audio/visual perception, within the framework of embodied simulation theory, will be reported. In Chapter 4, specifically, the creation of the custom-made video and music stimuli will be discussed.

4. Stimuli design

Two collaborations, with a filmmaker and film composer, lie at the heart of this interdisciplinary initiative between film studies, neuroscience, and the arts. The underlying aim of these collaborations was to create custom-made, naturalistic stimuli that would be ecologically valid but also carefully controlled for possible variables. The attention to detail and production value that was provided by this cross-disciplinary effort would not have been possible without the expertise of professional artists whom, in parallel, had a keen interest in the brain/bodily mechanisms underlying film and music.

In this chapter, I will discuss: 1) the model for the construction of the stimuli – the Soviet film *Letiat Zhuravli* (The Cranes Are Flying, Kalatozov, 1957), 2) the creation of the video stimuli with Fabrizio La Palombara and his team and 3) the creation of the musical stimuli with Eduardo Andrade.

4.1 *Letiat Zhuravli* (1957) as aesthetic model

Directed by Mikhail Kalatozov and filmed by Sergei Urusevsky, *Zhuravli* tells the story of two lovers, Veronika (played by Tatiana Samoilova) and Boris (played by Aleksei Batalov), whose blossoming romance is interrupted by the chaos and hardship of the Nazi Invasion. The film was an astonishing success both domestically and internationally. For Soviet viewers, it was a cathartic experience in which the War was finally told through unheroic, flawed, but deeply human characters (Shpolberg, 2017). Abroad, it stunned the West with its apparent "unsovietness" (Woll, 2003: 1). Bosley Crowther for the *New York Times* noted ironically, "Believe it or not, it is a picture about two young people romantically in love – in love with each other, that is, and not with a tractor or the Soviet state" (Crowther, 1960). Samoilova recalls receiving a gift at the East German film festival, a watch with the inscription: "Finally we see on the Soviet screen a face, not a mask" (Samoilova, 1993 cited in Woll, 2003: 77). The film was a transformative experience for an entire generation of future cineastes, from Andrei Konchalovsky ("It shook me") to Claude Lelouch ("After watching [*Zhuravli*] [...] I decided to become a director") (Letiat Zhuravli, 2016). In fact, the destinies of the French New Wave icon and *Zhuravli* were fatefully interlinked - it is thanks to Lelouch that *Zhuravli* garnered worldwide fame. At the time of its filming, the twenty-year-old Frenchman was permitted to travel to the USSR for two weeks. While in Moscow, filming clandestine footage of the city,

Lelouch manages to get access to Mosfilm on the very day *Zhuravli* was being filmed (specifically, the scene with Boris running up the stairs during his "death-fantasy"). Kalatozov invites Lelouch to stay along for the day to assist the shoot and visit the editing room. After seeing thirty minutes of rushes from the film, Lelouch is left "shaken and in tears". Upon his return to Paris, he immediately contacts the director of the Cannes Film Festival, informing him that the festival must get *Zhuravli* into the main competition at all costs (Claude Lelouch à propos, 2015). With Samoilova and Urusevsky as envoys, *Zhuravli* takes the 1958 Cannes Film Festival by storm, surpassing Bergman's *Brink of Life* and Tati's *Mon Oncle*, becoming the first and only Soviet film to receive the accolade of *Palme d'Or* as best film (Woll, 2003: 20).

The spectacular critical and box office success of *Zhuravli* is in large part due to Kalatozov and Urusevsky's commitment to innovatively and dynamically unleash the inner world of the characters through their "common understanding of movement as the essence of cinema" (Woll, 2003: 41). Alongside masters such as Vsevolod Pudovkin and Alexander Dovzhenko, Kalatozov owes much to Eisenstein, his teacher at VGIK. *Zhuravli*'s cinematic language, composed of long-shots, extreme low- and high-angle shots, hand-held camera, crane shots, expressive dollying and panning, and long-focus optics to convey expressiveness, sensitivity and emotional passion at a time when camerawork tended to be quite static, served to resurrect the forgotten energy of the Soviet silent era. Urusevsky's "subjective camera" garnered many avid but ultimately vacant efforts at technical imitation that failed to grasp the underlying expressive intention – to convey, through camerawork, Veronika's and Boris' inner human condition. Mitchell Lifton of the *Film Quaterly* writes:

> "What begins as a highly stylized visual convention suddenly turns into a very naturalistic one, and when we have become used to that, reverses itself again; the *adagio* of realism is reinvested with new meaning by the *allegro* of the montage scenes. What makes this apparent hodge-podge work is the fact that these movements, to extend the musical analogy, are both based on the same themes. A unity accrues which is not one of continuity but of rhythm, of the repetition and elaboration of certain symbolic elements, and of the way in which the camera treats these elements" (Lifton, 1960).

With the thaw of Krushev's de-stalinisation, the representation of Soviets was liberated from the confines of "civic virtues", and once movies could explore the "movement of the human spirit" anew, cinematography "regained its earlier dynamism" (Woll, 2003: 43). There are several memorable scenes in *Zhuravli* that seize the spectator with the dynamic "rhythm" of the camerawork and energetic montage. In the "death-fantasy" scene, Urusevky's camera circles around the birch trees as Boris begins to hallucinate the life he will never have, interwoven with an impressionistic montage of Boris running up the stairs and his imagined marriage to Veronika. Finalaly, fleeting slow-motion shots of Veronika in a wedding dress are superimposed on Boris' desolate surroundings as he dies. In the "suicide run" scene, Urusevsky's camera tracks Veronika as she desperately flees the hospital ward, ravaged by feelings of guilt for Boris' betrayal. The camera follows Veronika's suicide run to the train station, directly mirroring that of Tolstoy's *Anna Karenina*, before a chaotic montage of feet, fence, smoke and snow drive the rising tension as Veronika prepares to jump. During Veronica's run, Urusevsky this time put the camera directly into Samiolova's hands, who filmed herself on the run. Finally, there is the "staircase" scene (Fig. 4.1 – 4.8), at the opening of the film, before the War dismantles the bliss of Veronika's and Boris' romance. Returning to Veronika's apartment building shortly after dawn, the couple exchange soft whispers and delicate glances, with the camera framing them from a dramatic low angle. As Veronika darts up the stairs before the lovers decide when to see each other again, Boris hesitates for a few moments, calling Veronika, before darting up the stairs to follow her. The silence that enveloped the lovers' whispers is transformed into an aria of euphoric strings, as Boris sprints athletically up the stairs, occasionally looking up to spot Veronika, racing, dashing and lunging up the spiral staircase. Boris reaches Veronika, and the strings stop as abruptly as they started, as Boris asks "When?". Significantly, this scene is recalled in Boris' last dying moments, symbolizing both the euphoria of his love for Veronika before the coming of the war, his ascension to the heavens, and his continual searching for Veronika. A levy system was built in which a cradle harboring Urusevsky (Fig. 4.9) and his Konvas-Avtomat camera (4.11) were pulled up along an iron pillar as the cinematographer rose parallel to Boris, tracking him as he sprinted up the stairs, with operator rails ensuring that the camera cart followed a uniform, circular motion as it ascended the stairwell (Fig. 4.13) (Letiat Zhuravli, 2016). As with the rest of the film, the scene was as much an artistic achievement as a technological feat.

4.1 4.2

4.3 4.4

4.5 4.6

4.7 4.8

Fig. 4.1 – 4.8 | The "staircase" scene in *Letiat Zhuravli* (Kalatozov, 1957). Boris (Aleksei Batalov) runs up the stairs to his beloved Veronika (Tatiana Samoilova). Source: Kolesnikov, 2020g.

4.2 Video clips: recreating a classic scene

It was decided that the staircase scene would be used as a model for constructing the video stimuli of this project, due to its highly expressive use of cinematic movement to convey the pathos of joy through the symbolic motif of rising. According to Eisenstein, in order to evoke the strongest emotional impact on the spectator it is necessary to embody the emotional theme of a scene in its aesthetic composition. Kalatozov (a student of Eisenstein) and Urusevksy express the emotional theme of rising in all three aesthetic elements of the staircase scene: the actor, the music, and the camera.

An official University call was released for a director of photography (DoP) with knowledge of Soviet cinema and an interest in the applications of neuroscience research to film studies and filmmaking, with DoP Fabrizio La Palombara (Fig. 4.10) selected for the project. In order to achieve the complex camera movement in the scene and transform it into numerous, homogeneous clips to be used as experimental stimuli, it was decided together with our collaborator, the cinematographer Fabrizio La Palombara, that a drone would be used in order to attain the circular, fluid trajectory of the film without the lengthy and expensive process of building a levy. Drone pilot Sandro Russello was selected for this task, testing and selecting a drone model on location prior to filming. After conducting an inspection of over 20 potential film locations in Parma, the *Plesso della Pilotta* was chosen for its four-story tall rectangular stairwell, large enough to accommodate a drone, and a closed internal structure with a modern layout permitting safe access to electrical wiring (a permit was issued by the University Rector and the City of Parma). A casting call was carried out to select a male and female actor based on athleticism and similarity in height (it was decided that a female variation of Boris' sprint would be recreated for the purposes of controlling for gender). Subsequent to the consideration of over 100 applications, Leyla Tesanoviç and Alessandro Benini were selected for their physical agility and similarity in height (180 and 183 cm, respectively).

Several rehearsals were carried out with the actors and the drone pilot in order to fine-tune and synchronize the pace of the actors' sprints in relation to each other and the drone's flight path. Numerous tests were done to also verify that there was no electromagnetic interference

4.9

4.10

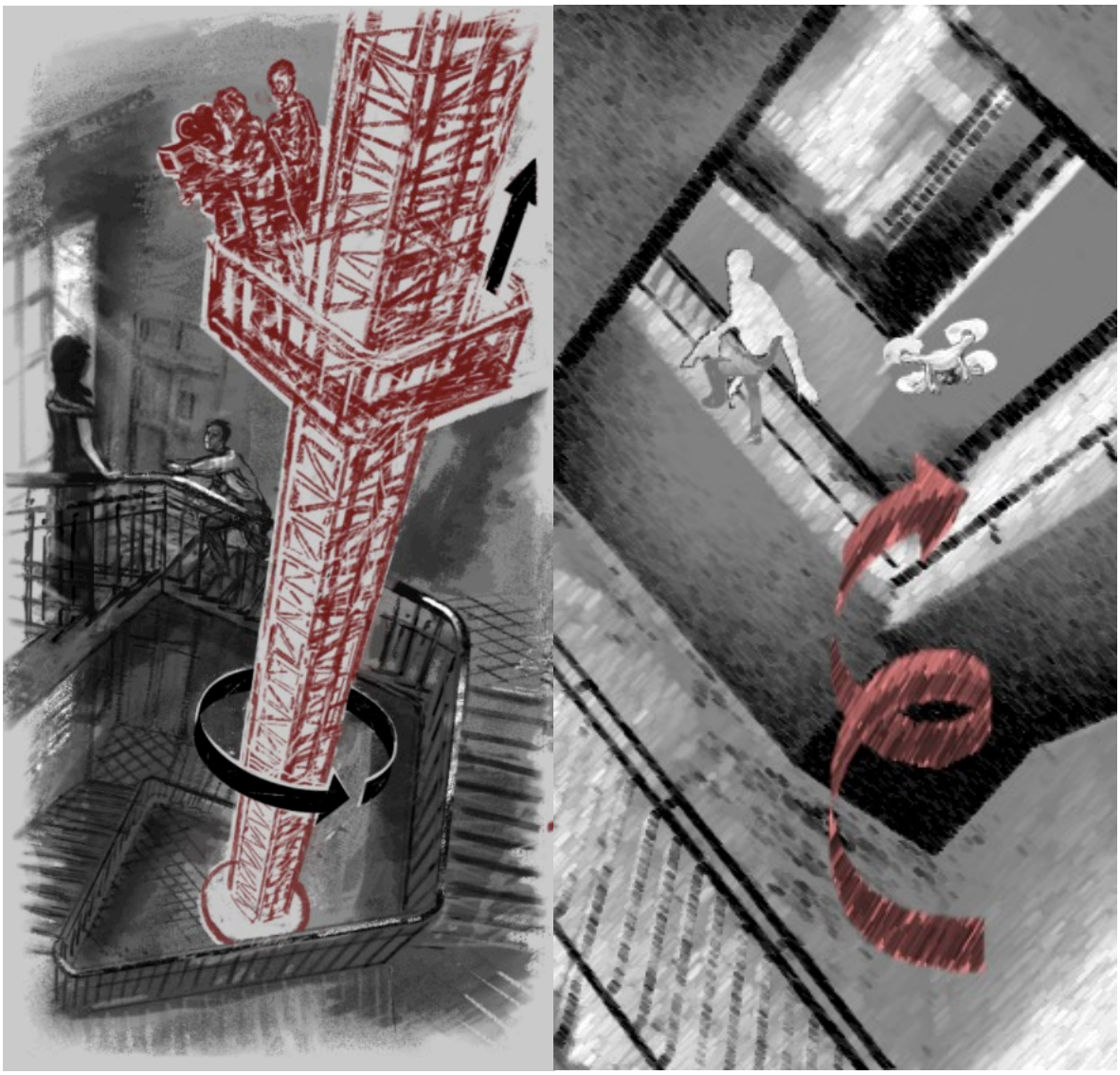
4.11

4.12

Fig. 4.9 – 4.12 | Cinematography of the "staircase" scene: then and now. Sergei Urusevsky, cinematographer on *Zhuravli* (1957), source: Global Cinematography Institute (4.9); Fabrizio La Palombara (4.10), cinematographer, designed the video shoot together with drone pilot Sandro Russello, source: *Fabrizio La Palombara*; Konvas (Konvas-Avtomat): portable 35mm motion-picture camera manufactured in the USSR (released in 1954) used to film *Zhuravli*, source: *Letiat Zhuravli*, 2016 (4.11); Drone model DJI Phantom 4 Pro Drone with camera model Sony Alpha 7S mounted on its platform, used to recreate the staricase scene in *Zhuravli* (4.12).

with the drone's trajectory and activity, and that the stairwell was large enough to ensure smooth and safe operating. On the day of shooting, lighting technicians rigged the location set with electric lights and taped down windows in order to control natural light. The actors performed ascending, descending and still renditions of the staircase scene from *Zhuravli*, with the DJI Phantom 4 Pro Drone (Fig 4.12) ascending and descending vertically up and down the stairwell while the Sony Alpha 7S mounted on its platform rotated 360 degrees on its axis (Fig. 4.14). The same variations were also filmed of the staircase without the actors.

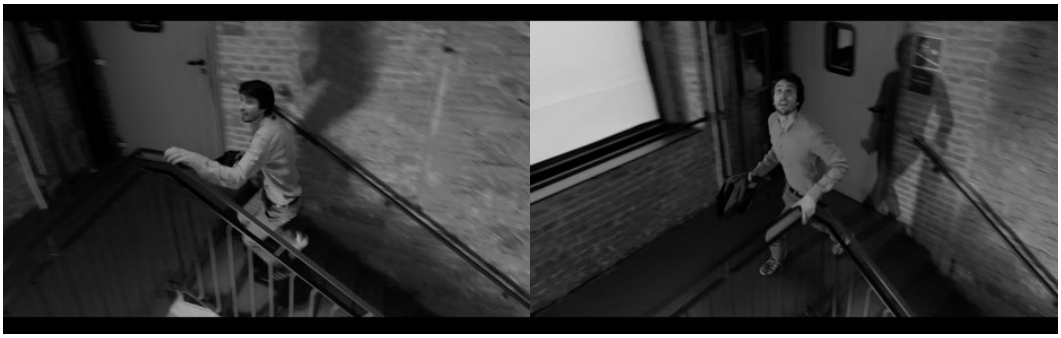
In this *hommage*, the objective was to recreate the dynamic fluidity of the camerawork and actor movement in *Zhuravli*'s staircase scene for the purposes of *empirically* investigating our sensorimotor engagement with film. To be sure, it was not possible to attain the artistic expressiveness and technological mastery of the original. However, this limitation was balanced out with the greater degree of control and manipulation of our variables of interest (the movement of camera/actor), with the possibility of recreating "ascending", "descending" and "still" variations in "male" (Fig. 4.15 – 4.22), "female" (Fig. 4.23 – 4.30) and "no actor" formats (Fig. 4.31 – 4.38). The next step was to compose the musical stimuli in order to, as with the video clips, optimize the manipulation of its embodied qualities.



4.13 4.14

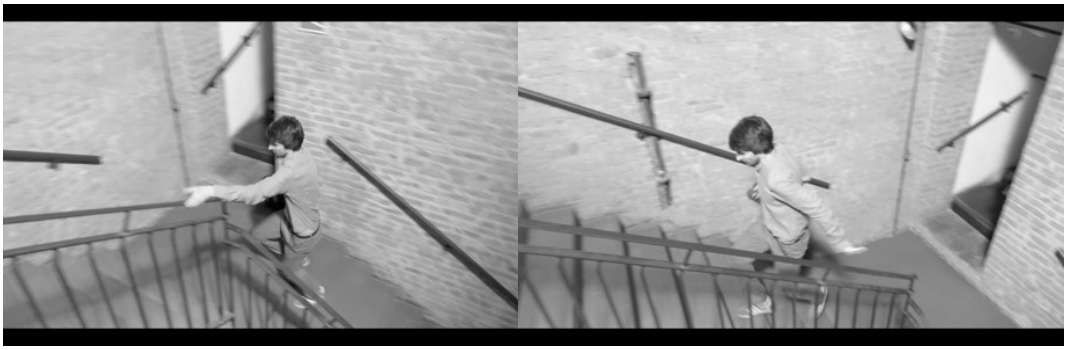
Fig. 4.13 – 4.14 | Rising while rotating. Sketch for filming of the "staircase scene" using a levy, a technical invention of cinematographer Sergei Urusevsky (4.13) (source: Letiat Zhuravli, 2016). Sketch of our recreation of the scene for filming using a drone (4.14).

4.15 4.16



4.17 4.18

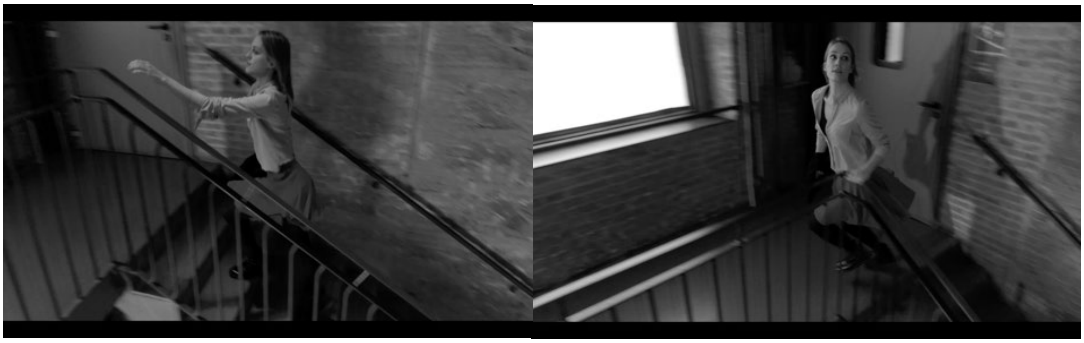
4.19 4.20



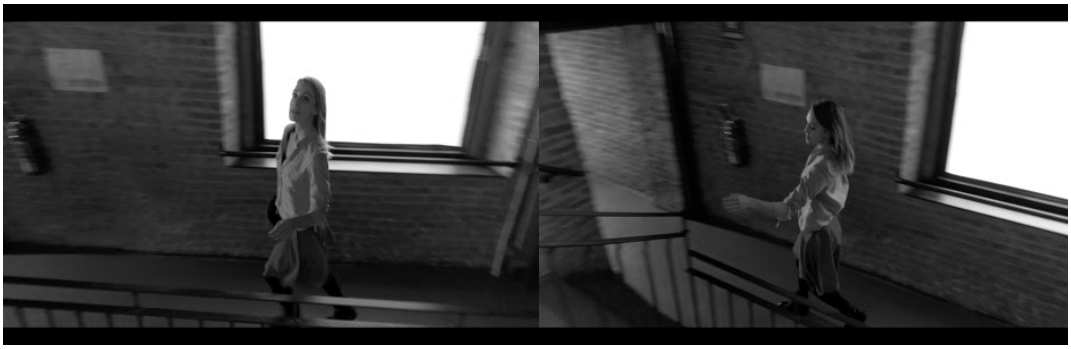
4.21 4.22

Fig. 4.15 – 4.22 | Recreation of the "staircase" scene: "Ascending Male" format. Actor: Alessandro Benini, film location: Plesso della Pilotta, Parma, Italy. Source: Kolesnikov, 2020a.

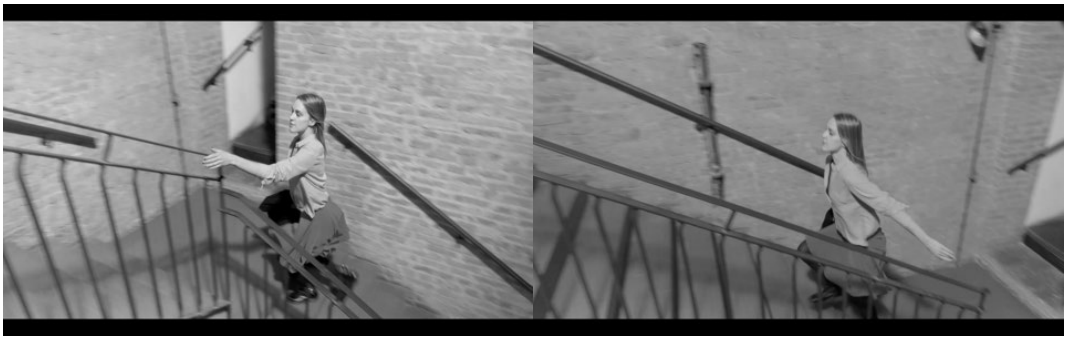
4.23 4.24



4.25 4.26



4.27 4.28



4.29 4.30

Fig. 4.23 – 4.30 | Recreation of the "staircase" scene: "Ascending Female" format. Actress: Leyla Tesanoviç, film location: Plesso della Pilotta, Parma, Italy. Source: Kolesnikov, 2020b.

4.31 4.32

4.33 4.34

4.35 4.36

4.37 4.38

Fig. 4.31 – 4.38 | Recreation of the "staircase" scene: "Ascending No Actor" format. Film location: Plesso della Pilotta, Parma, Italy. Source: Kolesnikov, 2020c.

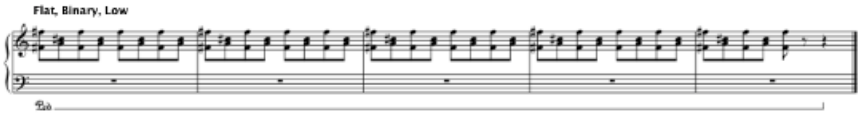
4.3 Musical tracks: composing movement

Though it is the visual vitality of the camerawork and actors' bodies that constitute the primary mode of embodied reception in *Zhuravli*, the music of Moisey Vaynberg (Fig. 4.39) is no less consequential to the emotional power of the film. The most memorial moments of the film, such as the staircase scene, the death fantasy and the suicide run, are preceded by a complete absence of non-diegetic sound, and are subsequently enveloped in Vaynberg's beautiful, powerful musical compositions. Without the expressive role of the music, Urusevsky's expertly filmed images lose their vitality – it is only through the convergence of the audio with the visual that the emotional power of the scene is consolidated. During Boris' death fantasy, the cries of Boris' fellow soldier Volodia are mixed with the joyful, ethereal strings of the staircase scene slowed down to match the slow motion of the wedding procession. Followed by the dramatic resounding of trumpets as the choral music accompanies images of Boris' and Veronika's imagined wedding, the last calls of the trumpet are heard as the drum roll quickens into a final crescendo and Boris falls to the ground. During Veronika's suicide run, a similar interplay between trumpet and strings accompany Veronika as she runs to the train station, the tension steadily rising in a crescendo before the sound of the approaching train drowns out the non-diegetic sound. Finally, in the short duration of the "staircase" scene (fifteen seconds), Vaynberg composes a track that accompanies Boris between the whispers he utters at the bottom and at the top of the staircase ("When?"). The artistic decision to include non-diegetic sound during Boris' sprint up the stairs is precisely what allocates the scene with emotional, symbolic and narrative significance. Vaynberg's composition embodies the motifs of joy and "rising toward love" in the scene – Boris' last happy memories, before the start of the war, are encapsulated in his upward ascent toward Veronika accompanied by Vaynberg's aforementioned high-pitched, joyful and ethereal strings.

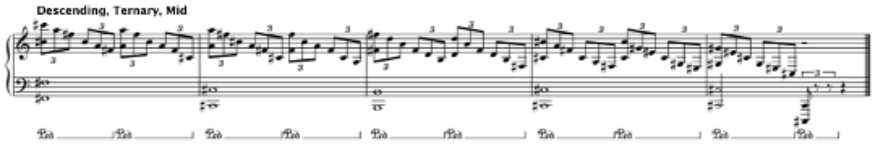
In a collaboration with film composer Eduardo Andrade (Fig. 4.40), a graduate of the Royal College of Music, musical variations to accompany the video clips were composed. The challenge lay in composing tracks that are effectively controlled for numerous acoustical features, such as duration, accent, timbre, tempo, and intensity, but that are still "music". In order to prioritize the investigation of perceived motion in melodic music, Andrade avoided the explicit expression of "valence" through key in his compositions,

4.39

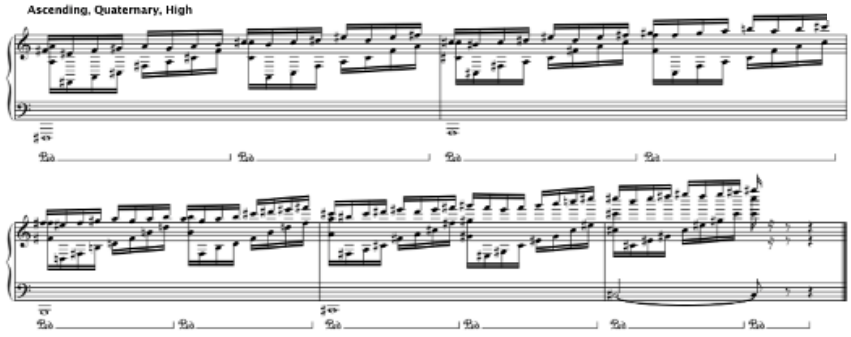
4.40



4.41



4.42



4.43

Fig. 4.39 – 4.43 | Composing movement. Moisey Vaynberg, composer of music for *Zhuravli* (1957), source: *Moisey Vaynberg* (4.39); due to restricted access, it was not possible to have access to the film score for *Zhuravli*; Eduardo Andrade, Royal College of Music, composer of the music stimuli used in this project, source: *Eduardo Andrade* (4.40); examples of tracks (labelling: Composition Contour, Note Pattern, and Composition Complexity: "Flat Binary Low" (4.41), "Descending Ternary Middle" (4.42), "Ascending Quaternary High" (4.43); the tempo is ♩ = 120, all excerpts and dynamics were played at *mf* as an average. Source: Kolesnikov, 2020d, 2020e, 2020f.

striving for emotional “neutrality”. Due to the greater ease of gradually integrating “Composition Complexity” through the addition of layers in the same instrument (low register notes for middle intensity and a higher melodic line for high intensity), and editing each individual recorded note's volume and duration, it was decided that Andrade would compose/produce the ten-second digital piano tracks using a MIDI keyboard to input the notes onto the Digital Audio Workstation (DAW) Cubase Pro. In order to provide the highest realism to the piano sound, the playback was configured to use Garritan/Abbey Road CFX Grand Piano samples. “Composition Contour” and “Note Pattern” were also integrated into the structure of the stimuli, and all notes were played with a certain degree of human variability in their dynamics. This approach was adopted by Andrade in order to increase likeness with the dynamics of how human interpretations are usually performed, as otherwise it would be evident that “computer programmed music” was being experienced, rather than “human performed music” (see Fig. 4.41 – 4.43 for examples and Appendix for complete partition of music stimuli).

5. Validation of video stimuli

5.1 Introduction

The unexpected discovery of mirror neurons at the turn of the 21st century revived the study of empathy and the link between action and perception. Giacomo Rizzolatti and colleagues isolated a population of neurons in the rostral part of the ventral premotor cortex (F5) of macaque monkeys which activate both during the execution of an action and when observing the action of a conspecific (Di Pellegrino et al., 1992; Rizzolatti et al., 1996; Gallese et al., 1996). It was observed that the same motor neurons are triggered when the monkey grabs a peanut, for example, as when the monkey observes another human or monkey perform the same action, leading to the automatic activation of the same neural mechanism which is triggered by the execution of the action. For the first time, a neural mechanism mapping the direct connection between the visual description of the characteristics of a motor act and its execution was identified (Gallese et al., 1996). Two other major experiments have demonstrate that mirror neuron activity reflects an internal, motor description of a perceived action rather than a purely visual description of its characteristics. In the first study, Evelyne Kohler and collaborators (2002) demonstrated that mirror neurons also encode the meaning of actions on the basis of sounds that correspond to them. A particular class of mirror neurons in the F5 area (audiovisual mirror neurons) respond not only when the monkey performs and observes a particular hand action, like breaking a peanut, but also when it hears the particular sound produced by the same action. In the second study, Maria Alessandra Umiltà and colleagues (2001) identified a subset of pre-motor mirror neurons that activate upon observing a partially concealed action, encoding the forthcoming action even in the absence of complete visual information. Thus, research on mirror neuron activity in macaques reveals the existence of a mechanism by which events perceived as different sounds or images are nevertheless coded as related to one another, insofar as they represent sensory features corresponding to the same motor act (Gallese, 2011).

Gallese and Alvin Goldman (1998) were the first to conceptualize mirror mechanisms in terms of simulation theory, later on spelled out as embodied simulation (ES) (Gallese, 2005, 2007, 2014). The term "embodied" signifies that body parts, bodily actions and body representations play a key role in cognition (Gallese and Sinigaglia, 2011). ES theory posits that humans tacitly

'simulate' the actions of the other by mapping them in the sensorimotor cortex of the brain. As mirror neurons anchor their multimodal processing in the neural mechanisms that govern our practical relationship to the world and others, social communication is enabled as the gap between self and others is reduced and we experientially understand the meaning of the actions performed by others (Gallese and Sinigaglia, 2011). It is hypothesized that the suppression of the sensorimotor alpha rhythm, also known as event-related desynchronization (ERD), during action execution and observation occurs as part of a neurophysiological mechanism known as motor resonance, or the mirror mechanism, facilitated by the activation of neurons resembling the mirror neurons discovered in the premotor and posterior parietal cortices of macaque monkeys. Indeed, a growing body of evidence corroborates the existence of a link between action execution and perception in humans (see Gallese, 2009).

It has been shown that numerous human cortical areas, such as the ventral part of the pre-central gyrus, the posterior part of the inferior frontal gyrus and the rostral part of the inferior parietal lobule are activated both during the execution and the observation of goal-orientated actions (see Frith and Frith, 2010; Hari and Kujala, 2009; Gallese, Keysers and Rizzolatti, 2004), such as videos of hand, mouth and foot actions (Buccino et al., 2001; Aziz-Zadeh et al., 2006). EEG studies have demonstrated a sensorimotor alpha ERD during the observation of different types of grasping (Muthukumaraswamy, Johnson and McNair, 2004), aimless movements (Babiloni et al., 2002) and videos of communicative hand gestures and grasping within social contexts (Streltsova et al., 2010). The observation of digital images of written language symbols (Heimann, Umiltà and Gallese, 2013) and abstract paintings (Umiltà et al., 2012; Sbriscia-Fioretti et al., 2013) have also been shown to modulate motor cortical activity. Furthermore, it has been demonstrated that the observer's motor repertoire or familiarity modulate motor cortical activity. In particular, Beatriz Calvo-Merino et al. (2005, 2006) and Guido Orgs et al. (2008) demonstrate that expert dancers observing videos of dancing steps in their particular professional style have stronger activation of motor areas with respect to when they observe dancing steps in which they do not have prior professional experience. A similar result was echoed in studies by Salvatore M. Aglioti et al. (2008) and Ana Maria Abreu et al. (2012) on motor activity during the observation of videos of basketball sequences by expert basketball players vs. nonathletes. This body of research supports the claim that the

link between action execution and perception constitutes the basis of social cognition, as concretized by embodied simulation theory, and the reported context sensitivity of visuomotor/sensorimotor activation supports the claim that these motor simulation processes are fine-tuned to serve specific social interactions.

It has been demonstrated that this corporal engagement occurs not only with others, but also with aesthetic phenomena (Freedberg and Gallese, 2007; Umiltà et al., 2012; Calbi et al., 2019; Siri et al., 2018, 2020; Ardizzi et al., 2020). In fact, the role of sensorimotor and visceromotor mechanisms in shaping film experience has been of interest to empiricists since the advent of cinema. In his recordings of the desynchronization of EEG rhythms in the alpha band (8-13Hz) from multiple electrodes, the filmologist Henri Gastaut and his collaborators were the first to demonstrate that motor cortex activation occurrs not only during action execution but also during the observation of actions; namely, the film footage of a boxer (Gastaut and Bert, 1954). The discoveries of mirror neurons and the direct link between perception and action, as well as embodied simulation theory, are informing a contemporary, interdisciplinary reassessment of how the brain-body system is engaged during film experience, in particular (Carluccio and Villa, 2006; D'Aloia and Eugeni, 2014; Coëgnarts and Kravanja, 2015; Carocci, 2018; Eugeni, 2018; Gallese and Guerra, 2012, see also 2020). According to the framework of *embodied cinema*, the meaning-making process in film is considered to be inextricably linked to the interrelation between the brain, body and environment of the viewer (Gallese and Guerra 2012, 2020). In their interdisciplinary collaboration, Gallese and Guerra in particular explore the types of empathic bodily engagement that exist in the art of the moving image, proposing a return to the brain-body in order to "capture our primordial contact with [...] film" and test the plausibility of certain theories circulating in film studies (Gallese and Guerra, 2012).

The film phenomenologist Vivian Sobchack identifies four basic types of movement in moving images. The first is the movement of human beings and objects within the frame; the second is the movement between images, namely editing; the third is the optical movement of the camera lens from a fixed position, such as zooming; and the fourth is the movement of the camera itself (Sobchack, 1982). Gallese and Guerra adopt a similar categorization in light of embodied simulation theory, which they believe can enrich the philosophical debate in film studies at the receptive and creative levels, shedding new light on at least three types of "film embodiment": 1) acting style, 2) film style, and 3) and the spectator's responses to filmed

bodies and objects (Gallese and Guerra, 2012: 198). The first stage of embodiment, the authors explain, is acting, which brings the audience into the forefront of "action" and "tactility", and film style (i.e., editing) emerges as a "negotiation" with the acting body. The role of the camera is integral here, endowing the cinematic experience with kinesthetic and tactile cues that animate the film with "vitalizing" qualities and a subjectivity of its own (Gallese and Guerra, 2014). Katrin Heimann and colleagues (2014) investigated the effect of camera movement on motor cortex activation in viewers, showing that the Steadicam elicits a stronger motor resonance compared to a "zoom". This is explained by the greater sense of "being there" that the Steadicam affords, in this way facilitating the motor mimicry of actions executed in the scene. The stronger motor resonance measured in the Steadicam condition may also be driven by motor engagement with the "trace" of the Steadicam's own movement across the scenic space. Replicating the study in an empty room, Heimann et al. (2019) found that greater motor resonance was again evoked for the Steadicam, providing the first empirical evidence that camera movement alone can modulate spectator's bodily engagement during film experience. In the domain of editing, Heimann et al. (2017) investigated sensorimotor engagement during the viewing of continuous and discontinuous cuts. Results suggest that "cutting across the line" (i.e., when the camera appears to 'jump' over an imaginary line between two actors) evokes a transitory spatial disorientation, leading to a violation of participants' sensorimotor expectations about the world, and a temporary suspension of embodied simulation of filmed actions. In *The Empathic Screen* (2020), Gallese and Guerra propose that cognitive neuroscience can make a valuable contribution to the study of human expression and aesthetic experience by shedding light on the role of the body in creative expression and reception, and in this way to "revitalize mimesis". However, no studies to date have interrogated the bodily basis of our engagement with film when camera movement, human movement and the perceived motion of music occur *simultaneously*, and the investigation of this multimodally embodied phenomenon is precisely the objective of this thesis.

As a first step, this validation will investigate the impact of camera and human movement on participant reception of appeal, involvement and perceived time. In light of the embodied attributes of slow motion and its ubiquitous use in cinema, its impact on the perception of movement on a behavioral level will also be investigated here. Indeed, little research has been done on how slowing the *speed* at which this movement is projected impacts participants' perception bodily involvement with the scene. "Slow motion" is commonly used in

audiovisual media to produce an array of effects that serve the narrative, the aesthetics, or purposes that are purely expressive (Rogers, 2013: 149). Whereas normal film speed is shot at 24 frames per second/Hz and projected with two to three repetitions of each frame to generate a smooth 28 or 72 Hz, slow motion images are shot at many more frames per second. When played back at a standard frame rate, time appears to slow down, and the screen reveals that which is ordinarily concealed to human perception. The slow-motion image disrupts the relation between moment and movement, and in the "extended moment" the "uncanny space" where movement approaches stasis is revealed (Sobchack, 2006). By virtue of the intimate relationship between cinematic technology and aesthetics, the slow-motion image immerses the spectator in micro-temporality and micro-movement that not only render the details of events such as explosions visible to the human eye, but also mimic extreme, life-threatening situations in which time is perceived to radically slow down. As revealed in studies on near-death experiences (Noyes and Kletti, 1976), sense of time is distorted by the activation of the norepinephrine system, which increases arousal and heightens the faculties of perception and attention, giving the impression of increased time to observe and react to one's surroundings. However, Clemens Wöllner and colleagues (2018) demonstrate that slow-motion scenes (i.e., commercial films, ballet and sports footage) actually evoke underestimations of duration, higher valence, and lower arousal compared to adapted real-time scenes. Another feature of slow motion is that it imitates the appearance of biological motion when carrying out high-precision tasks and responding to resistance.

In line with the growing trend of using naturalistic stimuli for cognitive science research on audiovisual media (Sonkusare, Breakspear and Guo, 2019), as explained in the previous chapter, an entirely new set of custom-made, ecologically valid video stimuli was created for the purpose of this project, modeled after the staircase scene from the *Zhuravli* (1957). Following an exegesis of Eisenstein's account of "mimetic contagion" in film, and in light of the predictions of embodied simulation theory and previous studies on sensorimotor engagement with camera movement, the aims of this validation are the following: 1) to investigate the impact of Camera Movement (Ascending, Descending, Still), Actor Presence (Female, Male, None [no actor]) and Image Speed (Normal, Slow, Very Slow) on participant ratings, and 2) to validate and select stimuli with the highest ratings for subsequent behavioral and EEG experiments.

The experimental hypotheses of this validation are the following:

1) Ascending conditions will have higher ratings for Movement, Emotional Involvement, and Physical Involvement than Descending and Still conditions (due to greater perceived exertion when the movement is upward, against the force of gravity)
2) Female and Male conditions will have higher ratings for Emotional Involvement and Liking than None (i.e., No Actor) conditions (due to increased identification when an actor is present in the video)
3) None conditions will have higher ratings for Physical Involvement than Female and Male conditions (due to the sensation that, when an actor is absent from the video, that the spectator *becomes* the drone).
4) Very Slow and Slow conditions will have higher ratings for Duration, Movement, Emotional Involvement and Physical Involvement with respect to Normal (due to greater perceived effort/exertion).

5.2 Materials and methods

5.2.1. Stimuli

Premiere Pro CC was used to edit the raw video footage into experimental stimuli in MP4 format with AAC and H.264 codecs and 1920 x 1080 resolution. Each clip had a frame rate of 25 frames per second, with a total of 250 frames or 10-seconds per clip. This duration was chosen with the future electroencephalography experiment in mind, which requires stimuli that are as short in length as possible. A cross dissolve of 25 frames (100ms) was included at the beginning and end of each clip in order to create a more fluid transition between the fixation cross and stimulus frames. The video clips were cut so that frame 26 coincided with the moment the actor's hand touched the banister (clips with no actor were cut as to visually coincide as much as possible with frame 26 in video clips with an actor) (see Fig. 5.1 – 5.9). In *hommage* to the original film, and in order to control for possible confounding effects, clips were converted into monochrome.

Fig. 5.1 – 5.9 | Video validation stimuli. Ascending conditions for Female, Male and None, respectively (5.1, 5.4, 5.7); Descending conditions for Female, Male and None, respectively (5.2, 5.5, 5.8); Still conditions for Female, Male and None, respectively (5.3, 5.6, 5.9).

Stimulus	Camera Movement	Actor Presence	Image Speed
1	Ascending	Female	Normal
2	Ascending	Male	Slow
3	Ascending	None	Very Slow
4	Ascending	Female	Normal
5	Ascending	Male	Slow
6	Ascending	None	Very Slow
7	Ascending	Female	Normal
8	Ascending	Male	Slow
9	Ascending	None	Very Slow
10	Ascending	Female	Normal
11	Ascending	Male	Slow
12	Ascending	None	Very Slow
13	Ascending	Female	Normal
14	Ascending	Male	Slow
15	Ascending	None	Very Slow
16	Ascending	Female	Normal
17	Ascending	Male	Slow
18	Ascending	None	Very Slow
19	Ascending	Female	Normal
20	Ascending	Male	Slow
21	Ascending	None	Very Slow
22	Ascending	Female	Normal
23	Ascending	Male	Slow
24	Ascending	None	Very Slow
25	Ascending	Female	Normal
26	Ascending	Male	Slow
27	Ascending	None	Very Slow
28	Descending	Female	Normal
29	Descending	Male	Slow
30	Descending	None	Very Slow
31	Descending	Female	Normal
32	Descending	Male	Slow
33	Descending	None	Very Slow
34	Descending	Female	Normal
35	Descending	Male	Slow
36	Descending	None	Very Slow
37	Descending	Female	Normal
38	Descending	Male	Slow
39	Descending	None	Very Slow
40	Descending	Female	Normal
41	Descending	Male	Slow
42	Descending	None	Very Slow
43	Descending	Female	Normal
44	Descending	Male	Slow
45	Descending	None	Very Slow
46	Descending	Female	Normal
47	Descending	Male	Slow
48	Descending	None	Very Slow

49	Descending	Female	Normal
50	Descending	Male	Slow
51	Descending	None	Very Slow
52	Descending	Female	Normal
53	Descending	Male	Slow
54	Descending	None	Very Slow
55	Still	Female	Normal
56	Still	Male	Slow
57	Still	None	Very Slow
58	Still	Female	Normal
59	Still	Male	Slow
60	Still	None	Very Slow
61	Still	Female	Normal
62	Still	Male	Slow
63	Still	None	Very Slow
64	Still	Female	Normal
65	Still	Male	Slow
66	Still	None	Very Slow
67	Still	Female	Normal
68	Still	Male	Slow
69	Still	None	Very Slow
70	Still	Female	Normal
71	Still	Male	Slow
72	Still	None	Very Slow
73	Still	Female	Normal
74	Still	Male	Slow
75	Still	None	Very Slow
76	Still	Female	Normal
77	Still	Male	Slow
78	Still	None	Very Slow
79	Still	Female	Normal
80	Still	Male	Slow
81	Still	None	Very Slow

Fig. 5.10 | Video validation stimuli /conditions. Total number of stimuli: 81.

A total of 27 clips were made with normal image speed (250 frames/clip). 27 clips were then slowed down to: 75% of the original image speed (333 frames/clip), and 27 clips were slowed down to 50% of the original image speed (125 frames/clip). In total, 81 video clips were created in each of the following modalities: Camera Movement (Ascending, Descending, Still), Actor Presence (Female, Male, None (i.e., No Actor)) and Image Speed (Normal, Slow (75%), Very Slow (50%)) (see Fig. 5.10). Luminosity analysis was performed on the Normal speed clips, in which it was verified that Ascending and Descending stimuli were significantly more luminous than Still stimuli. The stimuli set will be subsequently published as a paper and rendered open access for other researchers to use.

5.2.2. Participants

Participants were recruited using Facebook advertising that targeted our population of interest. 31 healthy volunteers of Italian nationality took part in the validation: 14 female and 17 males, mean age 25.03 (Standard Deviation – *SD* = 4.63, min = 18, max = 35). All participants had normal or corrected-to-normal visual acuity and were residents of Parma. Participants were asked to fill out an Audiovisual Survey (unvalidated) in order to screen for film/video making experience and the study of film at university level. We calculated the sample size by using G*Power 3.1 (Faul et al., 2007) and running a priori power analysis for a mixed-design multilevel linear model, with a level of power of 0.90, assuming a Cohen's F effect size equal to 0.25 (medium effect size), an alpha level of 0.05, and 5 measurements. All participants provided written informed consent to participate in the study, which was conducted in accordance with the Declaration of Helsinki (2013) and complied with the ethical standards of the Italian Board of Psychologists, as well as the Ethical Code for Psychological Research of Italian Psychological Society.

5.2.3. Procedure

Upon arrival, participants were asked to make themselves comfortable and were given information about the study. After giving informed consent, participants were seated approximately 60 cm from the screen of the computer, and stimuli were presented using a Windows PC. Once participants were comfortable, they were asked to fill out the following surveys: the Interpersonal Reactivity Index (IRI), the Immersive Tendencies Questionnaire (ITQ), the Visual Motor Imagination Questionnaire-2 (VMIQ-2), the Edinburgh Inventory,

and a brief survey in order to screen participants for film/video-making experience. Empathy was assessed in all participants using the IRI (Davis, 1980). The IRI consists of 28 self-report items, answered on a 5-point Likert scale ranging from 1 (“Does not describe me well”) to 5 (“Describes me very well”). The measure has 4 subscales, each made up of 7 different items that measure empathy as a multidimensional construct. The subscales of interest here are empathic concern (EC), which assesses "other-oriented" feelings of sympathy and concern for unfortunate others, and fantasy (F), which taps respondents' tendencies to transpose themselves imaginatively into the feelings and actions of fictitious characters in books, movies, and plays. Motor imagination was assessed in all participants using the VMIQ-2 (Roberts et al., 2008). The VMIQ-2 consists of 36 self-report items that measure the vividness of motor imagination, with three subscales: External Visual Imagination (i.e., imagining yourself carry out the movement as you observe from the outside), Internal Visual Imagination (i.e., imagining yourself carry out the movement through your own eyes), and Kinesthetic Imagination (i.e., imagining the physical sensation of carrying out the movement). Items are answered on a 5-point Likert scale ranging from 1 (“Perfectly clear and vivid, like in normal vision”) to 5 (“No image, only awareness of thinking about the movement”). Immersive tendencies were assessed in all participants using the ITQ (Witmer and Singer, 1998). The ITQ attempts to identify and measure possible individual differences in the abilities or tendencies of subjects to immerse themselves in different environmental situations. It uses a 7-point response scale that is based on the semantic differential principle, with each item being end-anchored by opposing descriptors, with the scale containing an anchor at the midpoint. The questionnaire consists of 34 self-report items, answered on a 7-point Likert scale. The subscale of interest here is the Involvement, which assesses the propensity to passively get involved in witnessing something, such as books, television and movies. Handedness was assessed in all participants using the Edinburgh Inventory (Oldfield, 1971). Containing 19 self-report items, the survey assessed whether participants were right-handed, left-handed or ambidextrous. Left-handed subjects were excluded. Half of participants were asked to complete the questionnaires before the computer task (see Fig. 5.11), and the other half afterwards. In the computer task, participants were presented with 81 video clips in randomized order, with each stimulus followed by 5 questions: 1) What long was the duration of the stimulus? 2) How much movement did you perceive?

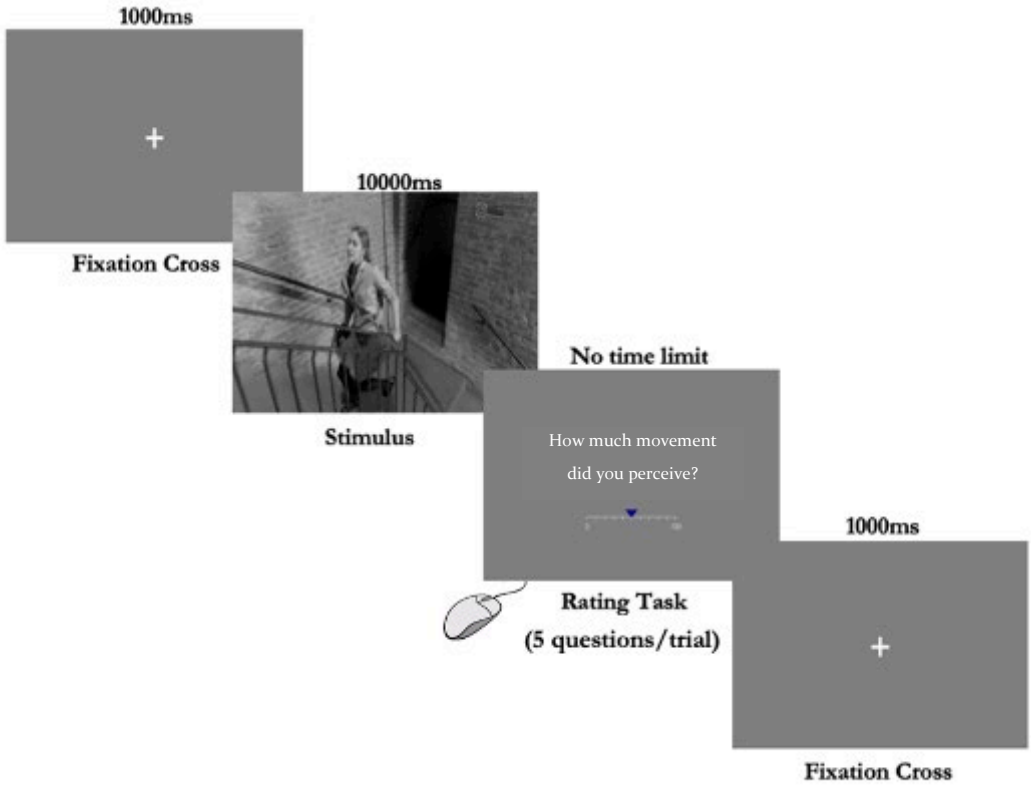


Fig. 5.11 | Example of video validation experimental trial. Components: fixation cross frame (1000ms), stimulus frame (10,000ms) and the Visual Analogue Scale (VAS) rating task (scale from 0 to 100, no time limit). Experiment was created using Psychopy 3.0.

3) How physically involved did you feel? 4) How emotionally involved did you feel? 5) How much did you like it? Participants were asked to observe the stimuli and answer the questions as quickly and accurately as possible (no specific time limit was given), using the mouse to click on the Visual Analogue Scale (VAS) ranging from 0 (very little) to 100 (very much). The study design was the following: 3 Camera Movements (Ascending, Descending, Still) * 3 Actor Presences (Female, Male, None) * 3 Image Speeds (Normal, Slow, Very Slow), for a total of 27 conditions. Each experimental condition was repeated three times, and with a total of 81 presented trials.

Before carrying out the computer task, participants were presented with a brief training phase to become accustomed with the procedure. After the experimental procedure participants were asked to fill out a short debriefing survey about their experience. The computer task was programmed using Psychopy 3.0 software (Peirce, 2011).

5.3. Results

In order to investigate whether VAS ratings are modulated by Camera Movement, Actor Presence and Image Speed, a linear mixed effect analysis was carried out. Participants' ratings were entered as dependent variables (Emotional Involvement, Physical Involvement, Movement, Liking, and Duration), Camera Movement (3 levels: Ascending, Descending, Still), Actor Presence (3 levels: Female, Male, None) and Image Speed (3 levels: Normal, Slow, Very Slow) as independent fixed variables, and participant intercepts as random effects. Tukey's test was used for post-hoc comparisons among means, where an error probability of less than 5% ($p < 0.05$) was considered significant in a population sample size of N=31. Correlations were performed between the total scores for the ITQ [Involvement], VMIQ2 [Internal], VMIQ2 [External], VMIQ2 [Kinesthetic], IRI [Empathic Concern], IRI [Fantasy], Audiovisual Survey and ratings during the experimental task (the critical probability values for multiple comparisons were corrected using the Bonferroni method (0.05/7 = 0.007)). All analyses were performed using R software (R Core Team, 2019) and lme4 (Bates et al. 2015), ordinal (Christensen, 2019), effects (Fox, 2003) and emmeans (Lenth, 2020) functions; for data visualization, the ggplot2 package was used (Wickham, 2016).

5.3.1. Physical Involvement

The model explained 66.53% of the variance in Physical Involvement ratings, taking into account the random effects (R^2_m =0.12, R^2_c = 0.67). The model revealed a significant main effect of Camera Movement ($\chi^2_{(2)}$ = 175.86, p < 0.0001), showing that participants felt more Physically Involved with Ascending than with Still (z = 12.14, SE = 1.26; Ascending: M = 40.6, SE = 3.53; Still: M = 25.3, SE = 3.53), and Descending more than Still (z = 10.69, SE = 1.26; Descending: M = 38.8, SE = 3.53), but no difference between Ascending and Descending (z = 1.45, SE = 1.26), see Figure 5.12. A significant main effect for Actor Presence was found ($\chi^2_{(2)}$ = 52.03, p <0.0001), showing that participants felt more Physically Involved with Female with respect to None (z = 5.79, SE = 1.26; Female: M = 37.0, SE = 3.53; None: M = 29.7, SE = 3.53), more with Male with respect to None (z = 6.62, SE = 1.26; Female: M =38.0, SE = 3.53), and with no significant difference between Male and Female (z = -0.83, SE = 1.26), see Figure 5.13. A significant main effect for Image Speed was found $\chi^2_{(2)}$ = 39.58, p < 0.0001), showing that participants felt more Physically Involved with Normal than with Slow (z = 4.19, SE = 1.26; Normal: M = 39.3, SE = 3.53; Slow: M = 34.0, SE = 3.53), and with Normal more than with Very Slow (z = 6.16, SE = 1.26; Very Slow: M = 31.5, SE = 3.52), see Figure 5.14. The model showed a significant Camera Movement*Actor Presence interaction ($\chi^2_{(4)}$ = 14.96, p < 0.01).

5.3.2. Emotional Involvement

The model explained 63.44% of the variance in Emotional Involvement ratings, taking into account the random effects (R^2_m = 0.075, R^2_c = 0.63). The model revealed a significant main effect of Camera Movement ($\chi^2_{(2)}$ = 13.74, p < 0.001), showing that participants found Ascending to be more Emotionally Involving than Descending (z = 3.48, SE = 1.26; Ascending: M = 37.6, SE = 3.43; Descending: M = 33.2, SE = 3.43), Ascending more than Still (z = 2.85, SE = 1.26; Still: M = 34.0, SE = 3.43), with no significant difference between Descending and Still (z =-0.63, SE = 1.26), see Figure 5.12. A significant main effect for Actor Presence was found ($\chi^2_{(2)}$ = 138.52, p < 0.0001), showing that participants found Female more Emotionally Involving than None (z = 10.99, SE = 1.26; Female: M = 40.3, SE = 3.43; None: M = 26.5, SE = 3.43), Male more Emotionally Involving than None (z = 9.15, SE = 1.26; None: M = 38.0, SE = 3.43), and no significant difference between Female and Male (z = 1.84, SE = 1.26), see Figure 5.13. A significant main effect for Image Speed was found $\chi^2_{(2)}$ = 6.44, p < 0.05), showing that participants felt more Emotionally Involved with Normal than with Slow (z = 2.12, SE = 1.26;

Normal: M = 39.3, SE = 3.43; Slow: M = 34.0, SE = 3.43), and with Normal more than with Very Slow (z = 2.27, SE = 1.26; Very Slow: M = 33.9, SE = 3.53), see Figure 5.14.

5.3.3. Movement

The model explained 72.93% of the variance in Movement ratings, taking into account the random effects (R^2_m = 0.40, R^2_c = 0.73). The model revealed a significant main effect of Camera Movement ($\chi^2_{(2)}$ = 1097.27, p < 0.0001), showing that participants perceived more Movement in Ascending than in Descending (z = 3.58, SE = 1.22; Ascending: M = 63.5, SE = 3; Descending: M = 59.1, SE = 3), Ascending more than Still (z = 30.31, SE = 1.22; Still: M = 26.4, SE = 3), and Descending more than Still (z = 26.73, SE = 1.22), see Figure 5.12. A significant main effect for Actor Presence was found ($\chi^2_{(2)}$ = 45.14, p < 0.0001), showing that participants perceived more Movement in Female than in None (z = 5.93, SE = 1.22; Female: M = 52.2, SE = 3; None: M = 44.9, SE = 3), and in Male with respect to None (z = 5.71, SE = 1.22; Male: M = 51.9, SE = 3), with no difference in Movement between Female and Male (z =0.22, SE = 1.22), see Figure 5.13. A significant main effect for Image Speed was found $\chi^2_{(2)}$ = 59.19, p < 0.0001), showing that participants perceived more Movement in Normal than in Slow (z = 4.66, SE = 1.22; Normal: M = 54.7, SE = 3; Slow: M = 49.0, SE = 3), in Normal more than Very Slow (z = 7.63, SE = 1.22; Very Slow: M = 45.4, SE = 3), and in Slow more than Very Slow (z = 2.98, SE = 1.22), see Figure 5.14. The model showed a significant Camera Movement*Actor Presence interaction ($\chi^2_{(4)}$ = 11.25, p < 0.0001).

5.3.4. Liking

The model explained 63.64 % of the variance in Liking ratings, taking into account the random effects (R^2_m = 0.06, R^2_c = 0.64). The model revealed a significant main effect of Camera Movement ($\chi^2_{(2)}$ = 28.68, p < 0.0001), showing that participants Liked Ascending more than Descending (z = 2.70, SE = 1.11; Ascending: M = 36.6, SE = 3.08; Descending: M = 33.6, SE = 3.08), Ascending more than Still (z = 5.36, SE = 1.11; Still: M = 30.6, SE = 3.08), and Descending more than Still (z = 2.70, SE = 1.11), see Figure 5.12. A significant main effect for Actor Presence was found ($\chi^2_{(2)}$ = 71.27, p < 0.0001), showing that participants Liked Female more than Male (z = 3.57, SE =1.11; Female: M = 38.1, SE = 3.08; Male: M = 34.1, SE = 3.08), Female more than None (z =8.41, SE =1.11; None: M = 28.7, SE = 3.08), and Male more than None (z = 4.84, SE = 1.11), see Figure 5.13. A significant main effect for Image Speed was found $\chi^2_{(2)}$ = 14.31, p < 0.001),

Camera Movement

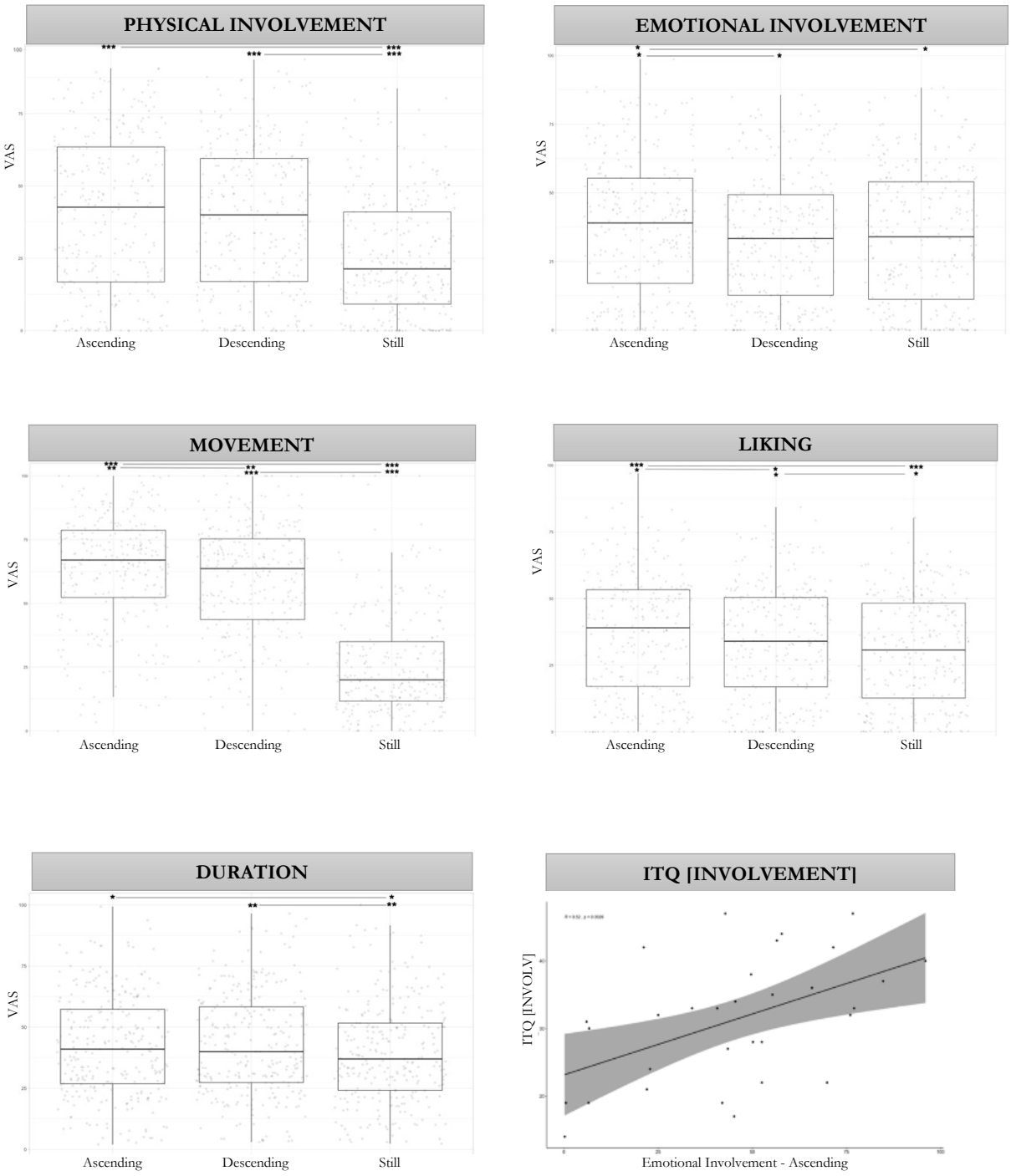


Fig. 5.12 | Boxplots and linear graph for main effect: Camera Movement. Boxplots depicting mean Visual Analogue Scale (VAS) ratings (on a scale from 0 to 100) for dependent variables "Physical Involvement", "Emotional Involvement", "Movement", "Liking" and "Duration" with respect to main effect "Camera Movement" (N=31). Linear graph depicting positive correlation between scores for Immersive Tendency Questionnaire (ITQ) [Involvement] (y-axis) and Emotional Involvement ratings for "Ascending" melodies (x-axis): p-value < 0.01, rho = 0.52 (N = 31). Error bars represent Standard Error (SE).

Actor Presence

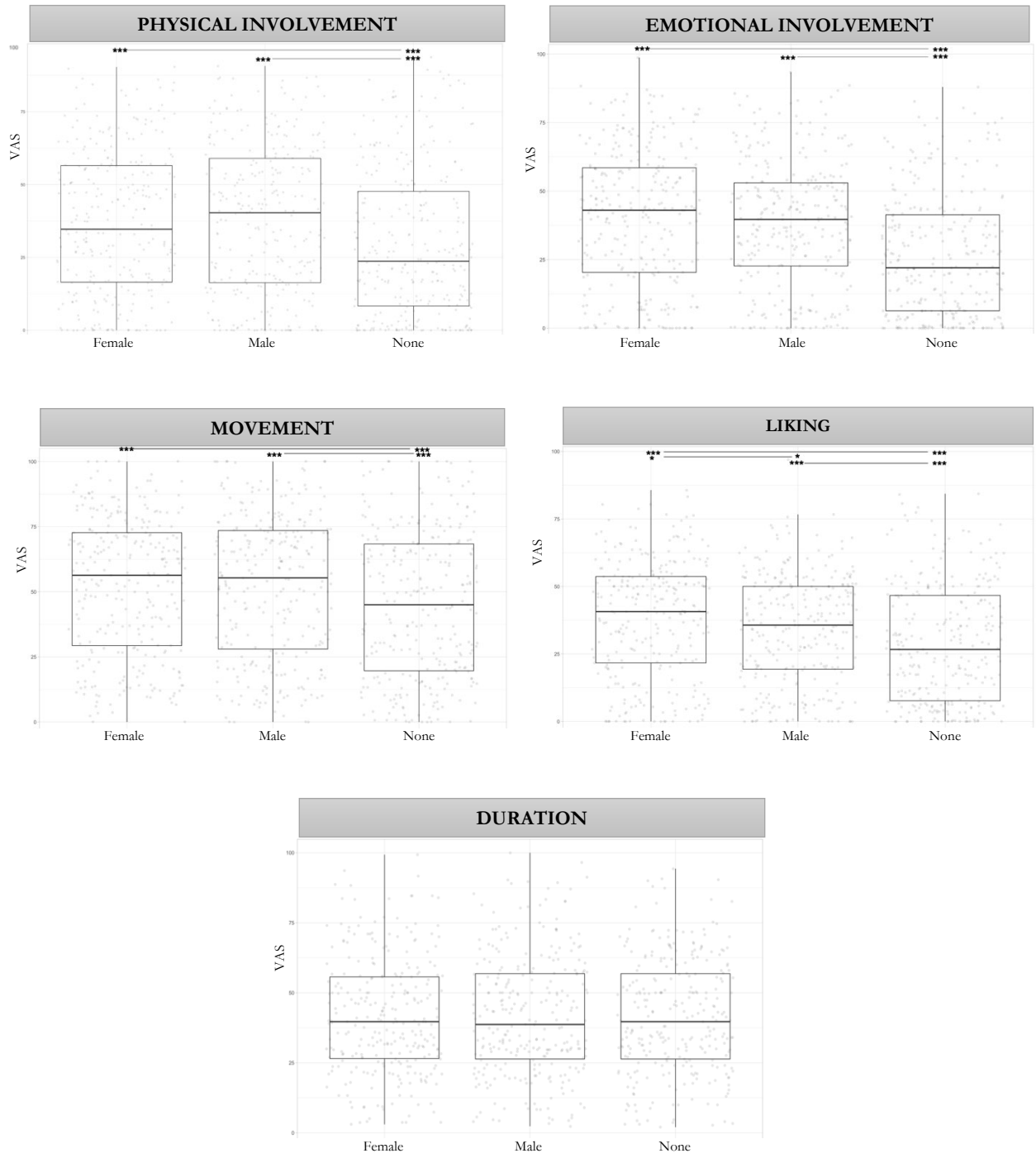


Fig. 5.13 | Boxplots for main effect: Actor Presence. Boxplots depicting mean Visual Analogue Scale (VAS) ratings (on a scale from 0 to 100) for dependent variables "Physical Involvement", "Emotional Involvement", "Movement", "Liking" and "Duration" with respect to main effect "Actor Presence" (N=31). Error bars represent Standard Error (SE).

Image Speed

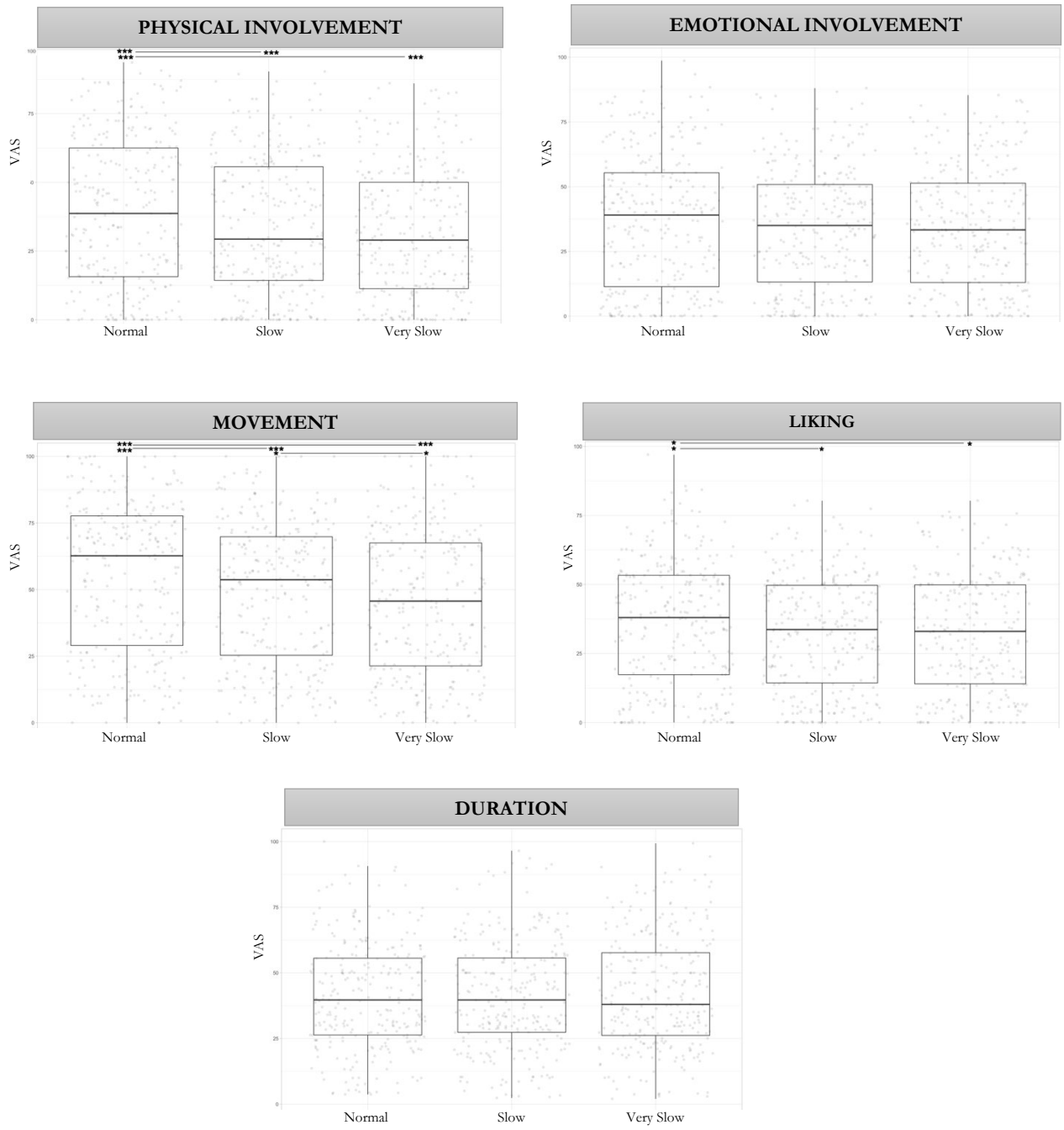


Fig. 5.14 | Boxplots for main effect: Image Speed. Boxplots depicting mean Visual Analogue Scale (VAS) ratings (on a scale from 0 to 100) for dependent variables "Physical Involvement", "Emotional Involvement", "Movement", "Liking" and "Duration" with respect to main effect "Image Speed" (N=31). Error bars represent Standard Error (SE).

showing that participants liked Normal more than Slow (z = 3.173, SE = 1.11; Normal: M = 36.0, SE = 3.08; Slow: M = 32.5, SE = 3.08), and Normal more than Very Slow (z = 3.371, SE = 1.11; Very Slow: M = 32.3, SE = 3.08), see Figure 5.14.

5.3.5. Duration

The model explained 72.02% of the variance in Liking ratings, taking into account the random effects (R^2_m = 0.01, R^2_c = 0.72). The model revealed a significant main effect of Camera Movement ($\chi^2_{(2)} = 15.07$, $p < 0.001$), showing that participants found Ascending to have a longer perceived Duration than Still (z = 2.95, SE = 0.95; Ascending: M = 42.2, SE = 3.27; Still: M = 39.4, SE = 3.27), and Descending more than Still (z = 3.66, SE = 0.95; Descending: M = 42.9, SE = 3.27, see Figure 5.12.

5.3.6. Questionnaire correlations

The [Involvement] items for the ITQ correlated with Emotional Involvement ratings for Ascending melodies (rho = 0.52, p-value < 0.01): the higher the ratings for ITQ [Involvement], the higher the Emotional Involvement ratings for Ascending melodies, see Figure 5.12.

5.4. Discussion

Previous studies have demonstrated that the observation of camera movement (Heimann et al., 2014, 2019) and human body movements and gestures in videos (Gastaut and Bert, 1954; Buccino et al., 2001; Aziz-Zadeh et al., 2006; Calvo-Merino et al., 2005, 2006; Orgs et al., 2008; Aglioti et al., 2008; Streltsova et al., 2010; Abreu et al., 2012) evoke sensorimotor resonance in the viewer. In the present study, we investigated the impact of Camera Movement (Ascending, Descending, Still), Actor Presence (Female, Male, None) and Image Speed (Normal, Slow, Very Slow) on participant ratings, in order to validate and select our stimuli for subsequent experiments. In order to extend the ecological validity of the study, we recreated the famously dynamic staircase scene of *Zhuravli* (1957), using the drone footage to edit a set of naturalistic stimuli. The task was to rate the stimuli for perceived Movement, Emotional and Physical Involvement, Liking, and Duration using a Visual Analogue Scale (VAS) ranging from 0 to 100, as quickly and as accurately as possible.

A significant main effect was found for Actor Presence. In particular, it was found that Female conditions had significantly higher ratings for Liking with respect to Male, but no significant difference was found between Female and Male for perceived Movement, Emotional Involvement or Physical Involvement. This may be explained by the fact that Female conditions were found to be more aesthetically pleasing in terms of appearance or movement fluidity. In line with our hypothesis, it was found that Female and Male had significantly higher ratings for Liking, Movement and Emotional Involvement with respect to None (No Actor) conditions. It was also found that None conditions were rated as significantly less Physically Involving than Female and Male conditions. This result provides support for embodied simulation theory, which suggests that we resonate more with conspecifics and familiar motor repertoires. Although the absence of others in the video image may increase the impression of "immersion" in the scene and, as a result, identification with the drone and simulation of its rotatory trajectory in space, participants still find the presence of an actor in the video image to be more involving on a motor level. Indeed, Female and Male conditions were rated as evoking significantly more Movement with respect to None conditions, which is in line with the "doubly" embodied nature of Female and Male conditions: whereas None conditions only have one vector of movement (i.e., the movement of the image as per the drone), Female and Male conditions have two (i.e., the drone and the actor's body).

A significant main effect was also found for Image Speed. In particular, it was found that, contrary to our hypothesis, no significant differences were found between conditions for perceived Duration. Thus, rather than a distortion in the perception of subjective time in slow motion conditions, participants perceived stimuli durations in terms of their actual running time, i.e., ten seconds for all stimuli. It was also found that Normal conditions had significantly higher ratings for Movement, Liking and Physical Involvement with respect to Slow and Very Slow conditions, contrary to our hypotheses. These results indicate that there was both greater preference and greater perceived embodiment for "ecological" movement (i.e., original image speed) than for "slow motion" movement. Normal Image Speed showed more movement in ten seconds than Slow and Very Slow image speed, and participants rated Movement precisely in terms of visualized apparent motion in a given time frame, rather than in terms of perceived effort. Seemingly, this appears to contradict the virtually ubiquitous praxis in the film industry of using slow motion in order to dramatize a scene and enhance

emotional involvement and intensity. However, the aforementioned results suggest that the emotional intensity of slow-motion scenes are actually largely driven by the presence of music, as found in a study by Wöllner and colleagues (2018), in which the presence of music was found to enhance the impact of stretched time, significantly augmenting physiological activation and perceived arousal and valence.

Finally, our results demonstrate that there is a significant main effect for Camera Movement. Specifically, it was demonstrated that, in line with our hypothesis, Ascending conditions have significantly higher ratings for Liking, Emotional Involvement and Movement with respect to Descending conditions. It was also found that, in line with our hypothesis, Ascending and Descending conditions have significantly higher ratings for perceived Movement, Physical Involvement and Liking with respect to Still conditions. Participants rated Ascending as more Emotionally Involving than Still, but, contrary to our hypothesis, no significant difference was found between Descending and Still. It was also found that the ITQ (Immersive Tendency Questionnaire) scores for the [Involvement] item correlated positively with Emotional Involvement ratings for Ascending melodies, specifically. No other significant correlations were found. Lastly, it was found that Ascending and Descending conditions were perceived to have a longer duration with respect to Still (although the actual stimuli durations were all ten seconds), supporting a study by Brown (1995) which demonstrates that stimulus motion lengthens perceived time.

The results for Camera Movement provide further evidence in support of embodied simulation theory, with results that are in line with prior neuroimaging studies on motor cortex activation during the observation of human and camera movement. Furthermore, a distinction has been demonstrated in the perception of Ascending movement with respect to Descending movement, in line with our hypotheses. There are several possible explanations for this result. Firstly, it may be that Ascending movement, which here is embodied in the act of running up the stairs, against the force of gravity, requires greater effort, or exertion, with this greater exertion being perceived as more movement, and more emotionally involving as a result. Neuroimaging studies are needed in order to investigate differences in motor activation between Ascending and Descending camera/actor movement. Another explanation for this result is that Ascending camera/actor movement is perceived as a representation of the motif of “rising”/ “upness”, and the embodied meaning of this gesture is felt to be more

emotionally involving with respect to “staying/ “standing” (Still) and “downward”/ “falling” (Descending). Both Ascending and Descending conditions portrayed a goal-orientated movement – going up toward something and going down toward something, but “ascent” felt more affecting than “descent” or “static”. Future studies could help elucidate how Ascending and Descending actor/camera movement are perceived in terms of valence, i.e., whether a more “positive” valence is attributed to Ascent and a more “negative” valence is attributed to Descent. Regardless, the greater Emotional Involvement of Ascending with respect to Descending conditions validates the artistic decision of Kalatozov in associating the joy of the protagonist to the motif of physical rising, epitomized by the staircase, as well as an affirmation to the emotional metaphor of joy/rising more generally.

6. Validation of musical stimuli

6.1 Introduction

Victor Zuckerkandl (1956) asserts that, "Whatever else music may be, one thing it must be: motion" (Zuckerkandl, 1956: 81). Although traditional cognitivist approaches consider music an auditory event, it is actually one of the most embodied experiences that belong to the repertoire of human activity. Continuing the work of musicologists such as Heinrich Schenker (1906/1954), Eduard Hanslick (1891), Carroll C. Pratt (1931/1968), Alexander Truslit (1938; see: Repp (1992) for translation), Roger Sessions (1941) and Zuckerkandl (1956) on the perceived "motion" of music, the philosophers George Lakoff and Mark Johnson (1988/2008) lay out key groundwork for understanding the perceptual structures of our experience of abstract musical phenomena in terms of the conceptual structures we use to describe music and musical meaning (Lakoff and Johnson, 2008). This relates closely to Stern's (2010) writings on the kinetic qualities, or vitality effects, of musical experience, in which embodied metaphors such as crescendo, decrescendo, accents, attack, staccato, legato, accelerando, and ritardando – all modes of movement – are so deeply encoded in our language that "it is almost impossible not to think of music in those terms" (Stern, 2010: 57). Johnson (2007) proposes that music is meaningful because it presents the "flow of human experience, feeling, and thinking in concrete embodied forms", and that this is meaning in the most profound sense (Johnson, 2007: 460). According to Margaret Walker, the "limitations of objective language" drive us to use metaphors, and "when metaphoric language fails, we are then driven to gestures" (Walker, 2000: 34). Gestures, in turn, express the elements of musical meaning that are out of the reach of verbal language (Godøy and Leman, 2010).

Metaphors are an effective mechanism for denoting musical meaning on a phenomenological, rather than associative, level, and the metaphorical constructions of language signify the myriad of bodily states that music triggers in us. Johnson writes, "Music does not re-present anything [...] the feeling is presented - enacted - in the felt experience of the listener. To hear the music is just to be moved and to feel in the precise way that is defined by the patterns of the musical motion" (Johnson, 2007: 461). Similarly, Susanne Langer explains that when a listener engages on an imaginative level with the development of temporal flow, pitch contour or intensity, the listener's experience has the "felt qualities of music" (Langer, 197:133).

In the domain of neuroscience, the role of sensorimotor engagement and simulation in musical experience is becoming a key area of interest in empirical studies of music, and the advent of neuroimaging techniques have made it possible to investigate the brain/bodily roots of embodied musical metaphors. A notable example of efforts to operationalize the notion of musical metaphor into a conceptual paradigm can be found in musicologist Arnie Cox's *mimetic hypothesis* (Cox 2001, 2011, 2017). Cox puts forward the important notion that musical meaning begins with the imitation of musical sounds and of the physical exertions that they produce. He separates the mimetic comprehension of music into two categories: mimetic motor action (MMA) as "overt" mimetic behavior, and mimetic motor imagery (MMI) for mimetic behavior that does not manifest in overt actions (i.e., "covert") (Cox, 2017). Thus, it is put forward that part of how we comprehend music is by imitating covertly (i.e., through thought, imagination and mental simulation, both voluntary and non-conscious) or overtly (i.e., muscle related responses such as head bobbing, foot tapping, air guitar, gesticulating) the observed sound-producing actions of performers (Cox, 2017: 12). When we do this, we are responding to two questions: *What is it like to do that?* and *What is it like to be that*?; the former describes the listener's desire to be the music, and the latter describes the response to an invitation to participate (Cox, 2017). Mimetic engagement refers to the more general notion that whenever we listen to music we are mimetically engaged, whether pre-reflectively or consciously. As auditory imagery includes mental representations of sounds that have been performed or may be performed, it is a kind of motor imagery. Infant studies demonstrate not only that as infants we imitate the facial expressions, vocalizations and gestures of those around us, but that infant-caregiver interactions are in fact mutual (Trevarthen and Malloch, 2000, 2017). Colwyn Trevarthen (1999) reports a video of a mother cradling and singing to a blind infant, in which the infant makes mimetic movements with her right arm in time with the mother's song as if conducting, stopping and resuming the movements when the song is interrupted or resumed.

A particular kind of motor imagery, related to planned, recalled or imagined speech actions, is known as *subvocalization,* or inner speech (Cox, 2017: 28). Subvocalization is a covert type of vocalization, such as the sound and feel of one's voice when reading, and *mimetic* subvocalization occurs when someone else's spoken words or singing is covertly imitated. A form of mimetic motor imagery, the mimetic hypothesis extends it to the imitation of musical sounds more generally (i.e., musical instruments). For most people the voice is our first and

most common means of auditory communication and how we understand the vocal expressions of others. Capable of imitating virtually all sounds without special training, the voice is "completely embodied within the flesh", and it is not limited to the larynx (Cox, 2017: 31). Head voice and chest voice refer to the sensation that the internal resonance of the voice seems to migrate from the chest to the head when one sings from low to high pitch, and it corresponds to the "pitch frequency is motion in vertical space" musical metaphor. Cox, in particular, refers to this phenomenon as the "human voice as source domain". Notably, pitch height is predominantly a Western concept, with other cultures using metaphors such as "young/old", "small/large" and "sharp/heavy" (Zbikowski, 2002). However, as Cox points out, medieval scholars such as Martianus and Isidore conceputalized vocalisations in terms of ascent (*ascendit*) and descent (*descendit*) centuries ago, with precedents in certain Ancient Greek and Hebraic writings (Cox, 2017: 86). Indeed, the visual representation of staff notation may originate as a metaphoric conceptualization of the greater or lesser exertion involved in singing, transcribed as greater or lesser "height". As Cox notes, the "greater is higher" metaphor can be observed in *Problemata*, where Aristotle writes "it is harder work to sing high pitches than low ones" (Aristotle, Problemata). Change of pitch in vocalization occurs through the tightening or relaxing of the vocal cords via the laryngeal muscles in coordination with muscles such as the diaphragm, which force air through the vocal folds (Cox, 2017: 90). As Sessions observes in *The Composer and His Message*:

> "If we instinctively respond to a rising melodic pitch by a feeling of increased tension and hence of heightened expression, or a falling pitch by the opposite sensation; if an increase in intensity of sound intensifies our dynamic response to the music, and vice versa, it is because we have already in our vocal experiences—the earliest and most primitive as well as later and more complicated ones—lived intimately through exactly the same effects. A raising of pitch or an increase in volume is the result of an intensification of effort, energy, and emotional power in the crying child just as truly as in the highly-evolved artistry of a Chaliapin or an Anderson" (Sessions, 1941: 108).

Indeed, from childhood we learn that if we raise our voice, we often end up raising pitch height – they go hand in hand. "Higher" pitch corresponds to greater tightening or muscular

tension, a feeling of greater *effort*, and thus a heightened affective state. For example, all singing involves activation of the abdominal muscles, shaping musical experience by way of correlations between abdominal exertions (tightness and relaxation) and emotional states. Drawing a similar example to Sessions, Cox describes the experience of listening to "Nessun dorma" from Puccini's "Turandot" as an example of the "greater [exertion] is higher" metaphor, in which one cannot help but mimic what it would be like to sustain the high B of the pathos-infused, final "Vincerò!": it feels like the ultimate "physical-artistic accomplishment" (Cox, 2017: 91). As the goal-orientated nature of music is often structured on climactic moments composed of melodic high points, and these highpoints often correlate with greater pitch height and greater acoustic strength (i.e., volume), the sense of "achievement" and emotional high-point in music is amplified.

There are numerous studies on the cross-modal interactions of music in terms of association with physical space and bodily motion, as well as evidence in support of voice-related subvocalization. In their theory of the co-representation of musical experience, Katie Avery and Istvan Molnar-Szakacs posit that the existence of a common neural substrate for music, language and motor functions is supported by evidence from studies of language disorders, in which children with dyslexia have been found to exhibit difficulties in musical timing (Overy et al., 2003), motor control (Wolff, 2002) and language (Goswami, 2015), as well as evidence of music lessons and speech therapy based on singing benefiting patients with severe non-fluent aphasia (Sparks et al., 1974). Robert J. Zatorre et al. (1996) demonstrate that perceiving a melody and imagining a melody involve overlapping neural systems, and activation of the supplementary motor area (SMA) occurs in both tasks. Cox points out that a theory of melodic recall interested solely in hearing would not predict activation of the SMA, which points to a combination of re-hearing and re-enacting, or re-singing. Gregory Hickok et al. (2003) instructed participants to listen to nonsense speech and music, followed by voluntary covert rehearsal, and found activation of shared auditory and motor brain areas in both perception and rehearsal. Consistent with the previous two studies, Daniel Callan et al. (2006) used fMRI in a similar procedure and found activation in the same and additional areas, in particular brain areas known to represent the lips and tongue. In order to exclude the possibility of linguistic processes, Zatorre and Andrea Halpern (1999) investigated auditory imagery for familiar melodies without lyrics, namely six-second portions of classical music, and film and TV themes. PET scans demonstrate activation of SMA in imagery for the

continuation of familiar melodies as well as in reimagining novel melodies, with the authors speculating that the SMA is involved in motor processes related to the formation auditory imagery regardless of familiarity. Studies have also demonstrated the existence of instrument-specific and rhythmic mimetic motor imagery. Halpern et al. (2004) instructed participants to make similarity judgements involving different kinds of musical instruments in a perceptual task and in an imagery task. The authors found that the SMA was activated, indicating a potential non-vocal motor imagery related to how the different instruments (e.g., a flute or trumpet) could be played. In a study by Simon Baumann and colleagues (2007), it was reported that during audio-motor coordination a network of predominantly secondary and higher order auditory and motor areas were activated in both musicians and non-musicians, with the lateral dorsal premotor cortex (dPMC) and the pre-supplementary motor cortex (preSMA) activation significantly increased for pianists. In two fMRI experiments, Joyce L. Chen, Penhune and Zatorre (2008a) had participants listen to stimuli played on a woodblock with the instruction that afterward they would be asked to tap along with the same rhythms. A separate group was then instructed to listen to the same stimuli without the instruction that they would be asked to tap along afterward. And yet, the same motor areas were activated in the two groups. Mid -premotor cortex (PMC) was significantly engaged not only during listening with anticipation and movement synchronization, but also during naive passive perception when there was no sound movement association. The dPMC was activated when subjects synchronized their movements, and the vPMC was activated when subjects listened with anticipation and synchronized their movements. The authors suggest that the vPMC maps a specific sound with a precise movement that produces that sound; thus, sounds must always be action-related for this region to be sensitive, and the dPMC implements the selection of movements based on higher-order rules such as those embedded in a rhythm's metrical structure. These findings corroborate a ventral-dorsal premotor dissociation, and related brain areas are thought to be involved in polymodal motion processing (Chen, Penhune and Zatorre, 2008b). Chen and colleagues posit that mid-PMC engagement may be a result of long-learned sound-movement associations, priming the brain for action. However, music perception activates the reticular formation in the brainstem, which is involved in responses such as the sudden release of energy to move during positive arousal. Reticular neurons, in particular, are responsible for the generation of motor reflexes through projections to spinal motoneurons, and Stefan Koelsch (2012) proposes that these projections

from the auditory brainstem to neurons of the reticular formation may contribute to the vitalizing effects of music and the human drive to move to music. It is highly likely that reticular formation activation in turn also influences the processing of new acoustic information (Koelsch, 2012: 95).

Though the human tendency to move in rhythmic synchrony with a musical beat is universal, it was thought until very recently that this faculty does not extend to non-human animals. In 2009, Aniruddh D. Patel and colleagues provided the first experimental evidence that non-human animals can spontaneously synchronize to a musical beat, demonstrating that the sulphur-crested cockatoo (*Cacatua gallerita eleonora*) is able to spontaneously adjust the tempo of its rhythmic movements to manipulations in tempo in a wide range of musical excerpts, and stay synchronized with the beat. According to the vocal learning and rhythmic synchronization hypothesis, entrainment to a musical beat relies on the neural circuitry for complex vocal learning, predicting that only vocal learning species such as humans and some birds, cetaceans and pinnipeds are capable of synchronizing movement to a musical beat (Patel et al., 2009). However, Andrew Rouse and co-authors (2016) demonstrate that the California sea lion (*Zalophus californianus*) is able to match head movements to isochronous repeating simple stimuli and adjust to novel tempos with introduced perturbations. These findings provide clear evidence for a shared faculty in beat keeping across multiple species, suggesting that the human ability to synchronize to musical rhythm stems from a broadly conserved, deeply rooted neurobiological mechanism.

While some cross-modal associations are innate, others may arise from learned, common environmental occurrences. James J. Gibson (1966, 1979) argues that an analysis of the environment is crucial to explaining behavior. Applying the Gibsonian perspective to the study of music, ecological acoustics is a discipline studying the relationship between human beings and their environment as mediated through sound, positing that listeners acquire practical knowledge from sounds in their acoustical surroundings rather than abstract concepts such as frequency, duration, and intensity. Niels Hansen and David Huron (2019) suggest that one such phenomenon is the perception of triplet rhythms. Although music scholars have long been aware of the sensation of rotation evoked in listeners, Hansen and Huron observe that no theory has been proposed to account for this “apparent association”. Drawing from the domain of ecological acoustics, the authors propose an Ecological Theory

of Rotating Sounds (EToRS) which maps fluctuations in loudness (i.e., accents) to trajectories of rotating sound sources. The authors ask the question, what kinds of rhythmic sounds are emitted when either a sound-producing object or listener is rotating? A recurring strong-weak pattern is most common for binary note patterns (two beats per measure), a strong-weak-medium or strong-medium-weak pattern for ternary note patterns (three beats per measure), a strong-medium-weak-medium pattern for quaternary note patterns (four beats per measure), and a strong-medium-weak-weak-medium pattern for quinary note patterns (five beats per measure). In the absence of an existing metrical context, the listener tends to hear the loudest event as marking the downbeat (Hansen and Huron 2019). Thus, according to the EToRS model, if the listener is situated outside the trajectory of the rotating sound, binary patterns reduce to pendular motion, whereas ternary reduce to spinning movement, and for quaternary and quinary note patterns it is ambiguous as to whether the motion is reduced to spinning or pendular motion. Testing their theory in a mixed population of professionals, semi-professionals and amateur musicians, the authors report that, 1) ternary note patterns are perceived as more spinning/rotating than non-ternary patterns, but that the difference is driven primarily by low rotating ratings for binary patterns, and 2) when pitch contour is moving, loudness patterns consistent with rotating trajectories are perceived as more rotating than unecological patterns, providing moderate support for an ecological account of rotating sounds. Given the rotatory quality of our video stimuli, by way of the rotatory trajectory of the drone during filming, it was deemed of interest to investigate the rotatory properties of music. Specifically, in addition to verifying the results of Hansen and Huron's study, constructing the musical stimuli set with "note pattern" as a variable of rotation would enable the subsequent investigation of perceived multimodal rotation in audiovisual stimuli.

Thus, as described in the aforementioned literature, there is considerable evidence and strong conceptual framework for the claim that an integral part of how we understand human movement and human-made sounds is in terms of our own experience of imitating the source of visual and auditory information, grounded in our own body's ability to make similar sounds. This evidence finds solid application with Eisenstein's writings on the melodic gesture, as explored in Chapter 2. Thus, the aims of the current study are to investigate music's "particular affinity" with movement in the following modalities: Composition Contour,

Composition Complexity and Note Pattern with respect to perceived Movement, Rotation, Direction, and Emotional and Physical Involvement.

The hypotheses are the following:

1) Ternary Note Pattern will be perceived as evoking more Rotation than Quaternary and Binary;

2) High Composition Complexity will be perceived as evoking for more Emotional Involvement, Physical Involvement and Movement with respect to Medium and Low; and

3) Ascending Composition Contour will be perceived as evoking more Emotional Involvement, Physical Involvement and Movement than Descending and Flat.

6.2. Materials and methods

6.2.1. Stimuli

The composer Eduardo Andrade created 27 ten-second piano tracks performed by him and edited/produced in a Digital Audio Workstation (DAW) Cubase Pro. To provide the highest realism to piano sound, the playback was configured to Garritan/Abbey Road CFX Grand Piano samples. The tracks were created in the following modalities: melodies with Ascending, Descending, or Flat Composition Contour, with Binary, Ternary or Quaternary Note Pattern, and with Low, Medium or High Composition Complexity (see Fig. 6.1). In total, 27 tracks were created. All tracks were balanced for volume and composition structure and formatted as waveform (WAV) audio files with a sample rate of 44.1 kHz and a bits per sample of 16. This set of stimuli will be subsequently published as a paper and rendered open access for other researchers to use.

6.2.2. Participants

Participants were recruited through opportunity sampling using Facebook, which filtered individuals for age, residence in Parma and a non-musician background. 30 healthy volunteers of Italian nationality took part in the validation: 16 females and 14 males, mean age 28.76 (*SD* = 44.64, min = 18, max = 35). All participants reported having normal hearing and normal or corrected-to-normal visual acuity. Participants were further screened for musical training prior to participation based on a screening questionnaire, in which individuals with

more than 6 years of music classes within the school curriculum or private musical tuition were excluded (including voice), with an average total number of years of musical training of 2 ± 1.75 SD years. We calculated the sample size by using G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) and running a priori power analysis for a mixed-design multilevel linear model, with a level of power of 0.90, assuming a Cohen's F effect size equal to 0.25 (medium effect size), an alpha level of 0.05, and 5 measurements. All participants provided written informed consent to participate in the study, which was conducted in accordance with the Declaration of Helsinki (2013) and complied with the ethical standards of the Italian Board of Psychologists, as well as the Ethical Code for Psychological Research of Italian Psychological Society.

6.2.3. Procedure

Before coming into the laboratory, participants were screened for musical training using a brief (unvalidated) survey on Google Forms, in order to ensure they met the criteria for 'non-musician'. Upon arrival, participants were asked to make themselves comfortable and were given information about the study. After giving informed consent, participants were seated approximately 60 cm from the screen of the computer and put on Sony WH1000XM2 Noise Cancelling headphones with an intensity at the sound source of 75% maximum volume. The stimuli were presented using a Windows PC and headphones. Once participants became comfortable, they were asked to fill out the IRI, the VMIQ-2, and the Edinburgh Inventory as in the Video Validation experiment. The Brief Musical Experience Questionnaire (BMEQ) was added to this study in order to assess participants for their musical experience in more detail. The BMEQ (Werner, Swope and Heide, 2006) consists of 53 self-report items, covering a wide range of topics pertaining to the place of music in a person's life, both for musicians and non-musicians. Two subscales were of interest here: affective reactions (A) (affective and spiritual reactions to music; 10 items) and reactive musical behavior (R) (motile reactions including humming and swaying, along with music; 9 items), with responses measured on a 5-point Likert scale ranging from 1 ("Very untrue") to 5 ("Very true"). Half of participants were asked to complete the questionnaires (using Google Forms) before the computer task, and the other half afterwards. In the computer task, participants were presented with 27 audio tracks in randomized order, with each stimulus followed by 1 question. There were 5 questions in total, and they were presented in randomized order: 1) How long was the duration of the stimulus?

Stimulus	Composition Contour	Note Pattern	Composition Complexity
1	Ascending	Binary	Low
2	Ascending	Binary	Medium
3	Ascending	Binary	High
4	Ascending	Ternary	Low
5	Ascending	Ternary	Medium
6	Ascending	Ternary	High
7	Ascending	Quaternary	Low
8	Ascending	Quaternary	Medium
9	Ascending	Quaternary	High
10	Descending	Binary	Low
11	Descending	Binary	Medium
12	Descending	Binary	High
13	Descending	Ternary	Low
14	Descending	Ternary	Medium
15	Descending	Ternary	High
16	Descending	Quaternary	Low
17	Descending	Quaternary	Medium
18	Descending	Quaternary	High
19	Flat	Binary	Low
20	Flat	Binary	Medium
21	Flat	Binary	High
22	Flat	Ternary	Low
23	Flat	Ternary	Medium
24	Flat	Ternary	High
25	Flat	Quaternary	Low
26	Flat	Quaternary	Medium
27	Flat	Quaternary	High

Fig. 6.1 |Table of music validation stimuli /conditions. Total number of stimuli: 27.

2) How much movement did you perceive? 3) How much rotation did you perceive? 4) How physically involved did you feel? and 5) How emotionally involved did you feel? Participants were asked to observe the stimuli and answer the questions as quickly and accurately as possible (no specific time limit was given), using the mouse to click on the Visual Analogue Scale (VAS) ranging from 0 (very little) to 100 (very much). For the "Duration" question, 0 signified "down" and 100 signified "up". The study design was the following: 3 Composition Contours (Ascending, Descending, Flat) * 3 Note Patterns (Binary, Ternary, Quaternary) * 3 Composition Complexities (High, Medium, Low), for a total of 27 conditions. Each experimental condition was repeated only once, with a total of 135 presented trials, see Fig. 6.2. Before carrying out the computer task, participants were presented with a brief training phase to become accustomed with the procedure. After the computer task, participants were asked to fill out a short debriefing survey about their experience. The experimental task was programmed using Psychopy 3.0 software (Peirce, 2011).

6.3. Results

In order to investigate whether VAS ratings are modulated by Composition Contour, Composition Complexity and Note Pattern, a linear mixed effect analysis was carried out. Participants' ratings were entered as dependent variables (Direction, Rotation, Movement, Emotional Involvement and Physical Involvement), and Composition Contour (3 levels: Ascending, Descending, Flat), Composition Complexity (3 levels: Low, Medium, High) and Note Pattern (3 levels: Binary, Ternary and Quaternary) as independent fixed variables, with participant intercepts as random effects. Tukey's test was used for post-hoc comparisons among means, where an error probability of less than 5% ($p < 0.05$) was considered significant in a population sample size of N=30. Correlations were performed between the total scores for VMIQ2 [Internal], VMIQ2 [External], VMIQ2 [Cinesthetic], IRI [Empathic Concern], IRI [Fantasy], BMEQ [A], BMEQ [R], and the Musical Survey and ratings during the experimental task (the critical probability values for multiple comparisons were corrected using the Bonferroni method (0.05/8 = 0.00625).

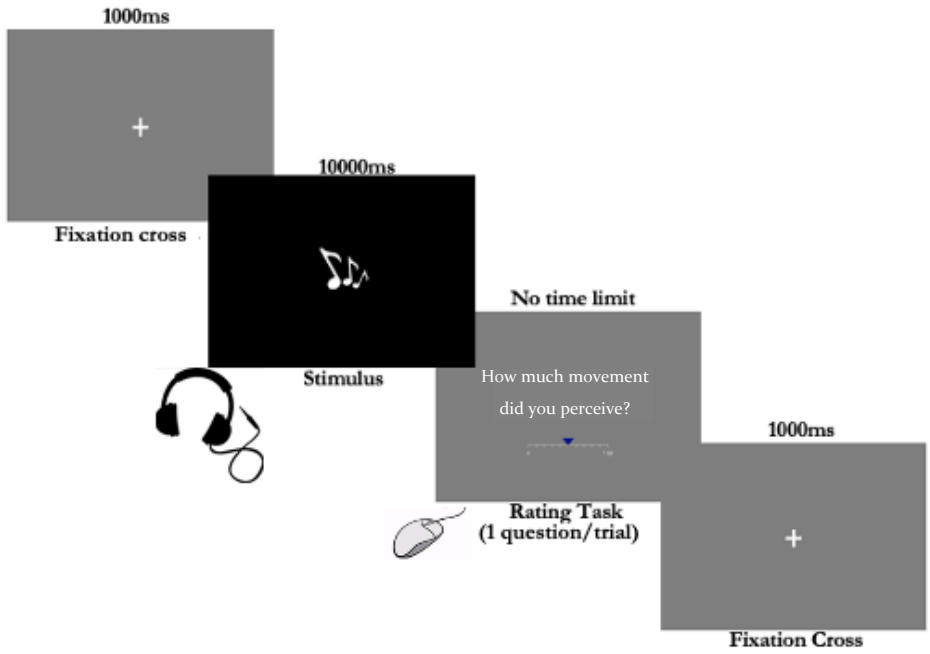


Fig. 6.2 | Example of music validation experimental trial. Components: fixation cross frame (1000ms), stimulus frame (10,000ms) and the Visual Analogue Scale (VAS) rating task (scale from 0 to 100, no time limit). Experiment was created using Psychopy 3.0.

All analyses were performed using R software (R Core Team, 2019) and lme4 (Bates et al., 2015), ordinal (Christensen, 2019), effects (Fox, 2003) and emmeans (Lenth, 2020) functions; for data visualization, the ggplot2 package was used (Wickham, 2016).

6.3.1. Direction

The model explained 55.76% of the variance in Direction ratings, taking into account the random effects (R^2_m = 0.5285211, R^2_c = 0.5576058). The model revealed a significant main effect of Composition Contour ($\chi^2_{(2)}$ = 869.6161, p < 0.0001), showing that participants perceived Ascending to be moving in a higher Direction with respect to Descending (z = 29.446, SE = 1.43; Ascending: M = 72.1, SE = 1.28; Descending: M = 30.0, SE = 1.28), Ascending to be moving in a higher Direction than Flat (z = 13.334, SE = 1.43; Flat: M = 53.0, SE = 1.28), and Descending in a lower Direction than Flat (z = -16.112, SE = 1.43), see Fig. 6.3. A significant main effect for Composition Complexity was found ($\chi^2_{(2)}$ = 40.7587, p < 0.0001), showing that participants perceived High Complexity melodies to be moving in a higher Direction than Medium (z = 5.989, SE = 1.43; High: M = 56.9, SE = 1.28; Medium: M = 48.3, SE = 1.28), and High moving in a higher Direction with respect to Low (z = 4.910, SE = 1.43; Low: M = 49.9, SE = 1.28), but no difference in Direction between Medium and Low (z =1.079, SE = 1.43), see Fig. 6.4. A significant main effect for Note Pattern was found $\chi^2_{(2)}$ = 39.8318, p < 0.0001), showing that participants perceived Quaternary to be moving in a higher Direction than Ternary (z = 3.308, SE = 1.43; Quaternary: M = 56.3, SE = 1.28; Ternary: M = 51.6, SE = 1.28), Binary moving in a lower Direction than Quaternary (z = -6.309, SE = 1.43; Binary: M = 47.3, SE = 1.28), and Binary moving in a lower Direction than Ternary (z = -3.001, SE = 1.43), see Fig. 6.5.

6.3.2. Rotation

The model explained 44.83 % of the variance in Rotation ratings, taking into account the random effects (R^2_m = 0.29, R^2_c = 0.45). The model revealed a significant main effect of Composition Contour ($\chi^2_{(2)}$ = 293.57, p <0.0001), showing that participants perceived more Rotation in Ascending than in Descending (z = 2.74, SE = 1.46; Ascending: M = 63.6, SE = 1.97; Descending: M = 59.6, SE = 1.97), more in Ascending than ini Flat (z = 16.02, SE = 1.46; Flat: M = 40.3, SE = 1.97), and Descending more than Flat (z = 13.28, SE = 1.46), see Fig. 6.3. A significant main

Composition Contour

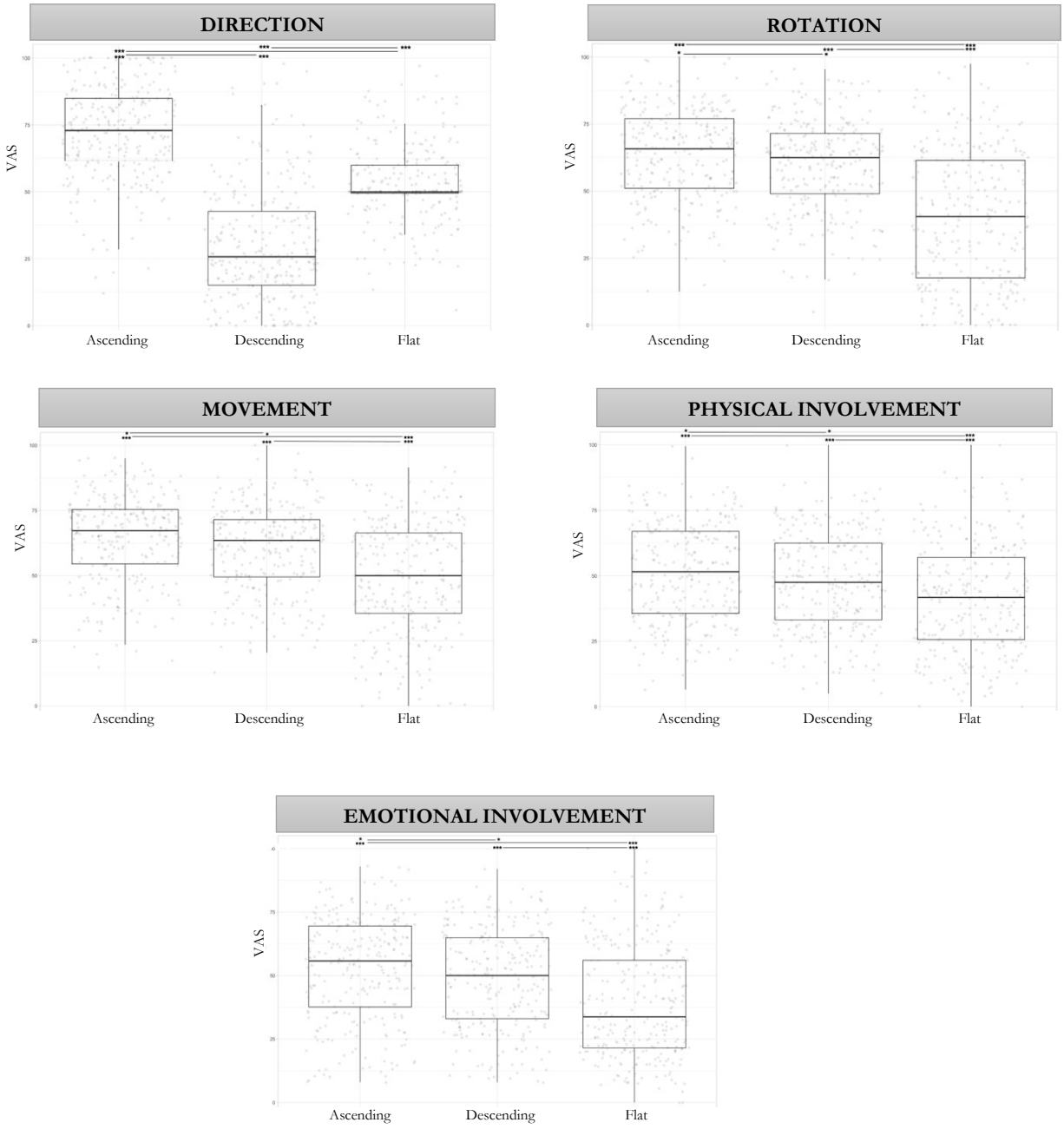


Fig. 6.3 | Boxplots for main effect: Composition Contour. Boxplots depicting mean Visual Analogue Scale (VAS) ratings (on a scale from 0 to 100) for dependent variables "Direction", "Rotation", "Movement", "Physical Involvement" and "Emotional Involvement" with respect to main effect "Composition Contour" (N=30). Error bars represent Standard Error (SE).

Composition Complexity

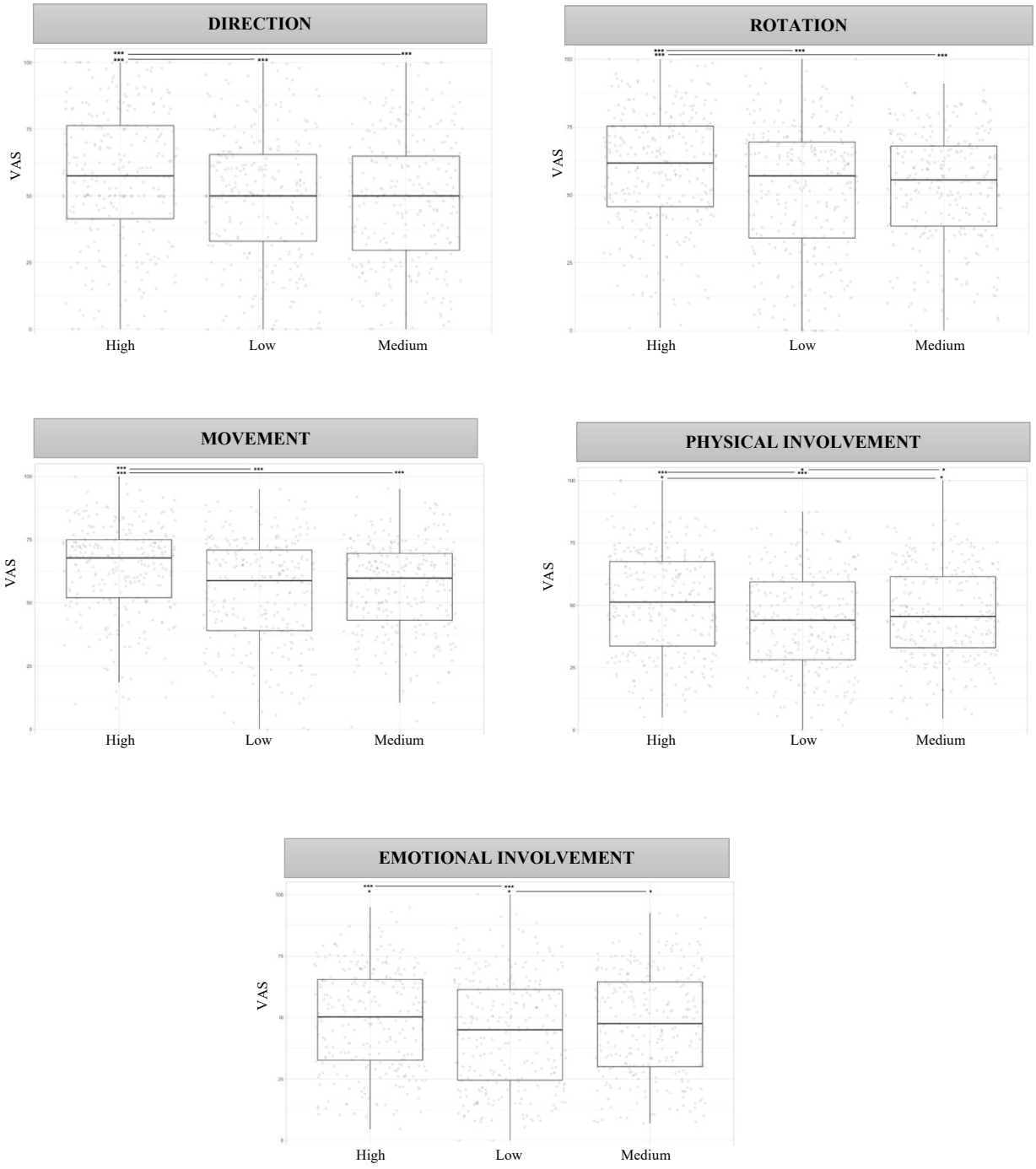


Fig. 6.4 | Boxplots for main effect: Composition Complexity. Boxplots depicting mean Visual Analalogue Scale (VAS) ratings (on a scale from 0 to 100) for dependent variables "Direction", "Rotation", "Movement", "Physical Involvement" and "Emotional Involvement" with respect to main effect "Composition Complexity" (N=30). Error bars represent Standard Error (SE).

Note Pattern

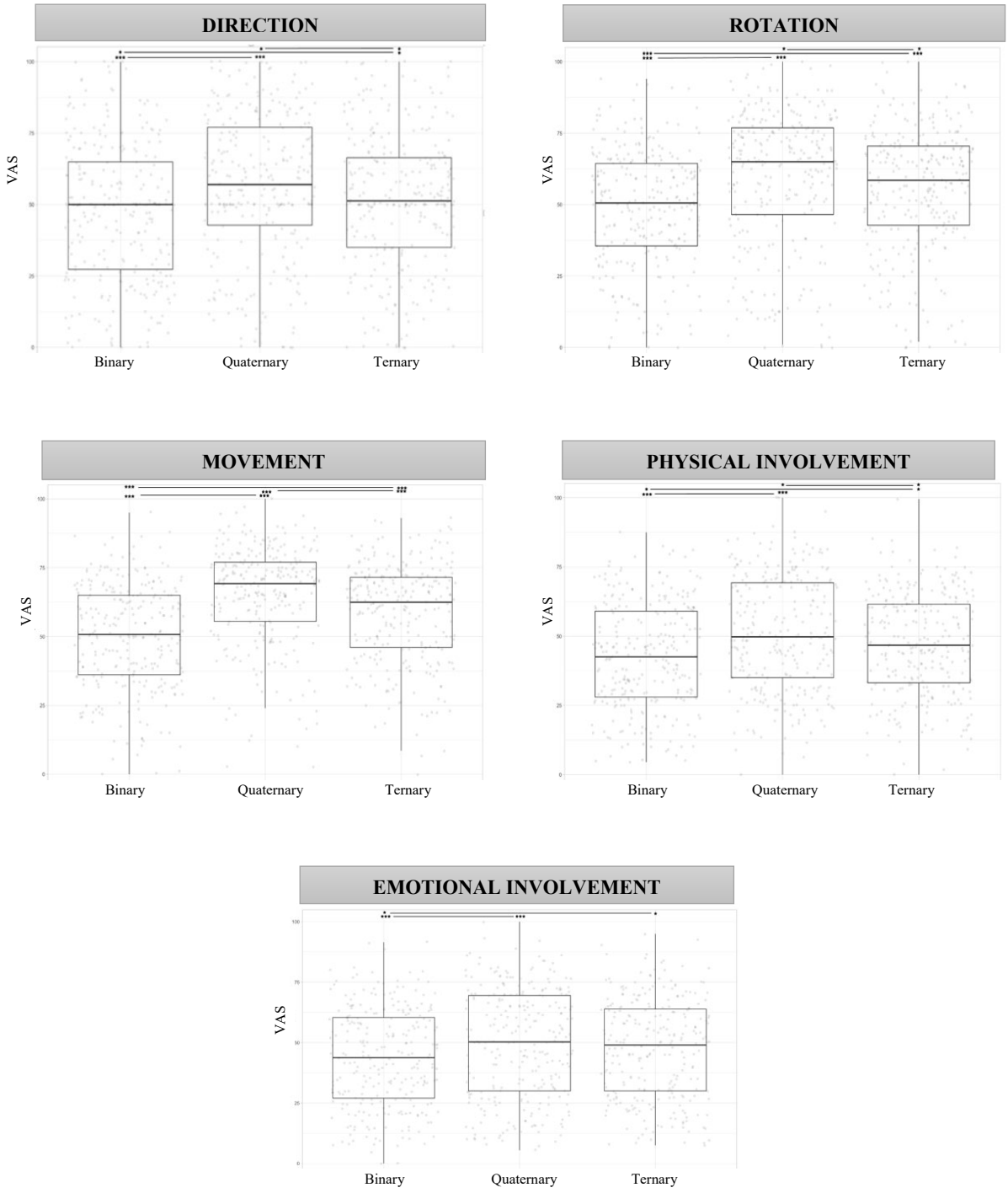


Fig. 6.5 | Boxplots for main effect: Note Pattern. Boxplots depicting mean Visual Analalogue Scale (VAS) ratings (on a scale from 0 to 100) for dependent variables "Direction", "Rotation", "Movement", "Physical Involvement" and "Emotional Involvement" with respect to main effect "Note Pattern" (N=30). Error bars represent Standard Error (SE).

effect for Composition Complexity was found ($\chi^2_{(2)}$ = 33.83, p<0.0001), showing that participants perceived more Rotation in High Complexity than Medium (z = 4.50, SE = 1.46; High: M = 59.4, SE = 1.97; Medium: M = 52.8, SE = 1.97), and in High with respect to Low (z = 5.44, SE = 1.46; Low: M = 51.4, SE = 1.97), but no difference in Rotation between Low and Medium (z = -0.94, SE = 1.46), see Fig. 6.4. A significant main effect for Note Pattern was found $\chi^2_{(2)}$ = 54.09, p < 0.0001), showing that participants perceived more Rotation in Quaternary than in Ternary (z = 2.51, SE = 1.46; Quaternary: M = 59.3, SE = 1.97; Ternary: M = 55.6, SE = 1.97), in Binary less than Quaternary (z = -7.24, SE = 1.46; Binary: M = 48.7, SE = 1.97), and in Binary less than in Ternary (z = -4.73, SE = 1.46), see Fig. 6.5. The model showed a significant Contour*Note Pattern interaction ($\chi^2_{(4)}$ = 12.56, p <0.01) and a significant Contour*Complexity interaction ($\chi^2_{(4)}$ = 21.15, p < 0.001).

6.3.3. Movement

The model explained 40.92% of the variance in Movement ratings, taking into account the random effects (R^2_m = 0.27, R^2_c = 0.41). The model revealed a significant main effect of Composition Contour ($\chi^2_{(2)}$ = 141.8837, p <0.0001), showing that participants perceived more Movement in Ascending than in Descending (z = 3.14, SE = 1.3; Ascending: M = 64.2, SE = 1.62; Descending: M = 60.1, SE = 1.62), Ascending more than Flat (z = 11.52, SE = 1.3; Flat: M = 49.2, SE = 1.62), and Descending more than Flat (z = 8.38, SE = 1.3), see Fig. 6.3. A significant main effect for Composition Complexity was found ($\chi^2_{(2)}$ = 55.16, p < 0.0001), showing that participants perceived more Movement in High than in Medium (z = 5.46, SE = 1.3; High: M = 63.3, SE = 1.62; Medium: M = 56.2, SE = 1.62), and more in High with respect to Low (z = 7.09, SE = 1.3; Low: M = 54.1, SE = 1.62), but no difference in Movement between Low and Medium (z =-1.63, SE = 1.3), see Fig. 6.4. A significant main effect for Note Pattern was found $\chi^2_{(2)}$ = 136.18, p < 0.0001), showing that participants perceived more Movement in Quaternary than in Ternary (z = 5.08, SE = 1.3; Quaternary: M = 65.1, SE = 1.62; Ternary: M = 58.5, SE = 1.62), in Binary less than in Quaternary (z = -11.64, SE = 1.3; Binary: M = 50.0, SE = 1.62), and in Binary less than in Ternary (z = -6.56, SE = 1.3), see Fig. 6.5. The model showed a significant Complexity*Note Pattern interaction ($\chi^2_{(4)}$ = 9.75, p < 0.05) and a significant Complexity*Contour interaction ($\chi^2_{(4)}$ = 15.37, p < 0.01).

6.3.4. Physical Involvement

The model explained 42.01 % of the variance in Physical Involvement ratings, taking into account the random effects (R^2_m = 0.12, R^2_c = 0.42). The model revealed a significant main effect of Composition Contour ($\chi^2_{(2)}$ = 61.05, p <0.0001), showing that participants felt more Physically Involved in Ascending than with Descending (z = 2.46, SE = 1.31; Ascending: M = 51.3, SE = 2.21; Descending: M = 48.1, SE = 2.21), Ascending more than with Flat (z = 7.65, SE = 1.31; Flat: M = 41.3, SE = 2.21), and Descending more than with Flat (z = 5.19, SE =1.31), see Fig. 6.3. A significant main effect for Composition Complexity was found ($\chi^2_{(2)}$ = 30.93, p < 0.0001), showing that participants felt more Physically Involved with High than with Medium (z = 3.03, SE = 1.31; High: M = 50.6, SE = 2.21; Medium: M = 46.7, SE = 2.21), more with High than with Low (z = 5.55, SE = 1.31; Low: M = 43.4, SE = 2.21), and less with Low than with Medium (z =-2.52, SE = 1.31), see Fig. 6.4. A significant main effect for Note Pattern was found ($\chi^2_{(2)}$ = 41.22, p < 0.0001), showing that participants felt more Physically Involved with Quaternary than with Ternary (z = 2.87, SE = 1.31; Quaternary: M = 50.9, SE = 1.31; Ternary: M = 47.2, SE = 2.21), with Binary less than with Quaternary (z = -6.41, SE = 1.31; Binary: M = 42.5, SE = 2.21), and with Binary less than with Ternary (z = -3.54, SE = 1.31), see Fig. 6.5. The model also showed a significant Complexity*Contour interaction ($\chi^2_{(4)}$ = 9.77, p <0.05).

6.3.5. Emotional Involvement

The model explained 53.45 % of the variance in Emotional Involvement ratings, taking into account the random effects (R^2_m = 0.11, R^2_c = 0.53). The model revealed a significant main effect of Composition Contour ($\chi^2_{(2)}$ = 120.95, p <0.0001), showing that participants felt more Emotionally Involved with Ascending than with Descending (z = 3.16, SE = 1.27; Ascending: M = 52.6, SE = 2.73; Descending: M = 48.6, SE = 2.73), with Ascending more than with Flat (z = 10.70, SE = 1.27; Flat: M = 39.0, SE = 2.73), and with Descending more than with Flat (z = 7.55, SE = 1.27), see Fig. 6.3. A significant main effect for Composition Complexity was found ($\chi^2_{(2)}$ = 22.35, p < 0.0001), showing that participants felt more Emotionally Involved with High Complexity than with Low (z = 4.64, SE = 1.27; High: M = 49.4, SE = 2.73 Low: M = 43.5, SE = 2.73), with Low less than with Medium (z = -3.12, SE = 1.27; Medium: M = 47.4, SE = 2.73), but no significant difference in Emotional Involvement between High and Medium (z =1.51, SE = 1.27), see Fig. 6.4. A significant main effect for Note Pattern was found $\chi^2_{(2)}$ = 21.30, p < 0.0001),

showing that participants felt less Emotionally Involved with Binary than with Quaternary (z = -4.55, SE = 1.27; Quaternary: M = 49.4, SE = 2.73; Binary: M = 43.6, SE = 2.73), and less with Binary than with Ternary (z = -2.93, SE = 1.27; Ternary: M = 47.3, SE = 2.73), but no significant difference in Emotional Involvement between Quaternary and Ternary (z = 1.62, SE = 1.27), see Fig. 6.5.

6.3.6. Correlations

No significant correlations were found between participants' ratings and questionnaire scores in this study.

6.4. Discussion

Previous studies have demonstrated that the perception of music evokes motor activation in human listeners (Zatorre et al., 1996; Halpern and Zatorre, 1999; Hickok et al., 2003; Baumann et al., 2007; Chen et al., 2008; Koelsch, 2013) and motor behavior in certain non-human animals (Patel et al., 2009; Rouse et al., 2016), indicating a phylogenetic association between music and movement. In the present study, we investigated the impact of Composition Contour (Ascending, Descending, Flat), Note Pattern (Binary, Ternary, Quaternary), and Composition Complexity (Low, Medium, High) on participant ratings, in order to validate and select our stimuli for subsequent experiments. In order to extend the ecological validity of the study, we collaborated with composer Eduardo Andrade to create 27 ten-second tracks recorded on a digital piano. The task was to rate the stimuli for perceived Direction, Movement, Emotional and Physical Involvement and Rotation using a Visual Analogue Scale (VAS) ranging from 0 to 100, as quickly and as accurately as possible.

Our results demonstrate a significant main effect for Composition Complexity. Specifically, in line with our hypothesis, High Complexity conditions are perceived as evoking more Movement, Rotation, Emotional and Physical Involvement and a higher Direction with respect to Medium and Low Complexity, with no significant difference found between Low and Medium Complexity for Direction, Rotation, Movement and Emotional Involvement. Low Complexity stimuli have the lowest note density (only middle high register was recorded), Middle Complexity have a higher note density (both middle high register and low

register were recorded), and High Complexity have the highest note density (both middle high register and low register were recorded, with an additional melodic layer added in the highest register, above the one in Low Complexity).These results suggest that notes are perceived as kinetic events, and the greater the note density, the greater the gestures are needed to produce the kinetic impressions, and thus the greater the perceived movement and involvement with the music.

Our results also demonstrate a significant main effect for Note Pattern. Specifically, contrary to our hypothesis, Quaternary conditions are perceived as evoking significantly more Rotation than Ternary and Binary conditions. It was also found that Quaternary conditions were perceived as having more Movement and being more Physically Involving with respect to Ternary and Binary conditions. Quaternary and Ternary conditions were found to be more Emotionally Involving than Binary conditions, but no significant difference was found between Quaternary and Ternary. Quaternary and Ternary conditions were found to be more Emotionally Involving than Binary conditions, but no significant difference was found between Quaternary and Ternary. Our results confirm Hansen and Huron's (2019) results that Binary patterns had the lowest ratings for rotation, but we found that participants perceived significantly more Rotation in Quaternary patterns (not in Ternary patterns). A lack of control for the confounding effects of melodic accent structure may explain this result. However, it was necessary to control for the confounding effects of tempo, which is fastest in Quaternary note patterns due to their higher note density (4 beats per measure instead of 3 or 2) but equivalent duration (all stimuli are ten seconds long). Indeed, our results are in line with Hansen and Huron's results that perceived rotation increases with tempo. It was also demonstrated that participants perceive more Rotation in Ascending melodies than in Descending melodies, and more in Ascending and Descending melodies than in Flat melodies, in line with Hansen and Huron's result that the sensation of rotation is perceived in loudness patterns consistent with rotating trajectories in pitch that is moving.

Finally, our results demonstrate that there is a significant main effect for Composition Contour. For the Direction rating, participants were instructed that ratings approaching "0" signify "downward" and ratings approaching "100" signify "upward", with "50" signifying "no change in Direction". It was demonstrated, in line with our hypothesis, that Ascending

conditions are perceived as moving upward (M = 72.1), Descending conditions are perceived as moving downward (M = 30.0), and Flat conditions are perceived as moving neither upward nor downward (M = 53.0). This result confirms that participants are able to identify the melodic contour of the melodies; however, it is not possible to conclude whether participants actually perceive the changing "direction" of the melodic contour, or whether they judge the direction of the melody based on the last set of high or low notes, which are easier to discern. More studies are needed to confirm this result. Of particular interest is the result that participants rate Ascending melodies as evoking more Movement and Emotional and Physical Involvement than Descending Melodies. It was also demonstrated that Ascending and Descending melodies are rated as having significantly more Movement and being more Emotionally and Physically Involving than Flat conditions. These results provide novel evidence in support of the mimetic subvocalization theory of Cox's embodied music cognition framework. This result indicates two important things. Firstly, even though the tempo, interval size and note density are equivalent in all conditions, participants not only perceive more Movement in Ascending and Descending conditions with respect to Flat, but more Movement in Ascending than in Descending conditions. This suggests that participants associate change in Movement and Physical Involvement with change in Composition Contour, and thus that participants associate Ascending Composition Contour with greater exertion or effort, in line with Cox's Mimetic Motor Imagery (MMI) and Voice as Source Domain hypotheses (i.e., the "greater [exertion] is higher" metaphor). In addition to the greater effort perceived in the high, harder-to-reach notes of Ascending melodies, participants may also associate Ascending piano melodies with the felt vibration of sung high notes in the head region, i.e., head voice (as opposed to the felt vibration of sung low notes in the chest, i.e., chest voice).

Secondly, results provide preliminary evidence for increased Emotional Involvement in response to Ascending melodies (with respect to Descending and Flat melodies), suggesting a possible link between perceived exertion, mimetic participation and musical affect. As Cox explains with the high B of the final "Vincerò", the goal-orientated nature of tonal music commonly features high points (Eitan, 1997), often correlating with greater pitch height and greater acoustic strength (volume), thereby enhancing the sense of "physical-artistic" accomplishment (Cox, 2017). However, the volume across all audio stimuli in the study were

kept constant, suggesting that Ascending Composition Contour is perceived as more Emotionally Involving than Descending Composition Contour irrespective of acoustic strength. These results also mirror the Camera Movement results of the Video Validation, in which Ascending Camera Movement was found to evoke more Movement and be more Emotionally and Physically Involving than Descending or Still Camera Movement, suggesting the existence of a cross-modal mapping of Ascending movement, exertion and affect.

Altogether, these results support the embodied music cognition framework (Godøy and Leman, 2010; Cox, 2017) and the mimetic subvocalization hypothesis (Cox, 2017), in particular. They also provide empirical evidence for enduring notions of melodic motion (e.g., Hanslick, 1891; Schenker, 1906/1954; Truslit, 1938; Pratt, 1931/1968; Sessions, 1941; Zuckerkandl, 1956; Johnson, 2007). Future studies are needed to investigate the perception of musical gestures in other timbres (such as strings and woodwind), the perception of rotation in melodies with different note patterns that are controlled for note density (i.e., by modified accent structures) and in association with rotating video modalities, and the impact of musical gestures of contour and rotation on sensorimotor activation.

7. Behavioral (audiovisual) study and plan for electroencephalography (EEG) study

As outlined in the literature reviews of the Video and Music Validations, the grounding of visual and musical experience in embodied formats is supported by both strong conceptual frameworks and empirical evidence from neuroscientific studies. Furthermore, results from these two Validations demonstrate that Ascending videos and Ascending melodies are found to be more Emotionally and Physically Involving and perceived to have more Movement not only more than Still videos and Flat melodies, respectively, but also more than Descending videos and Descending melodies, in support of the theories of embodied simulation and mimetic subvocalization, in which it is predicted that we simulate the effort needed to carry out a particular action.

In this study, the notion of the audiovisual gesture, as anchored in Eisenstein's writings, will be operationalized as part of a novel investigation into "unimodal" and "multimodal" gestures. Following the Validations of the video and music stimuli, the aims of the current study are to investigate: 1) how actor/camera movement and melodic movement occurring individually (i.e., Audio or Video Modalities) affect participant ratings (in the behavioral study) and desynchronization of the *mu* rhythm (in the EEG study) with respect to actor/camera and melodic movement occurring simultaneously (i.e., Audiovisual Modality), and 2) how the presence of Contour movement (i.e., Ascending conditions) affect participant ratings and *mu* rhythm desynchronization with respect to absence of Contour movement (i.e., Still). Such an investigation will enable future studies on how correspondence between multimodal gestures affect participant ratings and motor resonance. Due to the ongoing COVID-19 pandemic, however, the EEG investigation had to be postponed until safe conditions could be guaranteed for participants and researchers.

7.1. Behavioral materials and methods

The hypotheses for the behavioral study are the following: 1) Ascending conditions will have significantly higher ratings for perceived Movement and Physical and Emotional Involvement with respect to Still conditions, and 2) Audiovisual conditions will have significantly higher

ratings for perceived Movement and Physical and Emotional Involvement with respect to Audio and Video conditions.

7.1.1. Stimuli

Validated audio and video stimuli were selected to create audiovisual stimuli for this experiment (see Fig. 7.1). Still and Ascending video clips at Normal Image Speed in Female and Male format (MP4, H.264 codec, 25 frames/second) were selected from the Video Validation, as results show that there are no significant differences for Movement and Physical and Emotional Involvement ratings between the two genders, and Ascending clips have higher ratings than Descending clips (a total of 6 video clips). Quaternary audio tracks (WAV, 44.1 kHz sample rate, 16 bits per sample) were selected from the Music Validation as results show that they have the highest Movement and Physical and Emotional Involvement ratings with no significant differences between Low, Medium and High Composition Complexity (a total of 6 audio tracks), and like in the Video Validation, Ascending conditions have higher ratings than Descending conditions. Audio and Video stimuli were combined using IMovie in order to export new audiovisual stimuli as MOV files (1920 x 1080, Timecode, Apple ProRes 422, Linear PCM codecs) that do not cause audio data compression. In total, 18 audiovisual monochrome stimuli were created in each of the following modalities: Contour (Ascending, Still) and Modality (Audio, Video, Audiovisual) (see Fig. 7.2 – 7.4 for "Audiovisual Still" stimuli and Fig. 7.5 – 7.7 for "Audiovisual Ascending" stimuli).

7.1.2 Participants

Participants were recruited through opportunity sampling using Facebook, which filtered individuals for age, residence in Parma and a non-musician and non-filmmaker background. 34 healthy volunteers of Italian nationality took part in the validation: 17 females and 17 males, mean age 26.06 (*SD* = 4.99, min = 18, max = 39). All participants reported having normal hearing and normal or corrected-to-normal visual acuity. Participants were further screened for musical training prior to participation based on a screening questionnaire, in which individuals with more than 6 years of music classes within the school curriculum or private musical tuition (including voice) were excluded. Participants were also screened for prior filmmaking experience based on a screening questionnaire, with no professionals or film scholars recruited. We calculated the sample size by using G*Power 3.1 (Faul, Erdfelder, Lang,

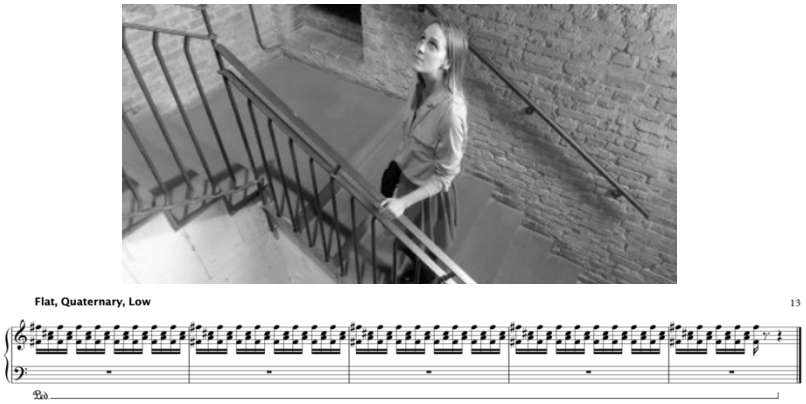
& Buchner, 2007) and running a priori power analysis for a mixed-design multilevel linear model, with a level of power of 0.90, assuming a Cohen's F effect size equal to 0.25 (medium effect size), an alpha level of 0.05, and 3 measurements. All participants provided written informed consent to participate in the study, which was conducted in accordance with the Declaration of Helsinki (2013).

7.1.3. Procedure

Before coming into the laboratory, participants were screened for musical training using a brief (unvalidated) survey on Google Forms, in order to ensure they met the criteria for "non-musician". Upon arrival, participants were asked to make themselves comfortable and were given information about the study. After giving informed consent, participants were seated approximately 60 cm from the screen of the computer and put on Sony WH1000XM2 Noise Cancelling headphones with an intensity at the sound source of 75% maximum volume. The stimuli were presented using a Windows PC and headphones. Once participants became comfortable, they were asked to fill out the IRI, the VMIQ-2, and the Edinburgh Inventory, as in the Video Validation. Half of participants were asked to complete the questionnaires before the computer task using Google Forms, and the other half afterwards. In the computer task (see Fig. 7.8), participants were presented with 18 audiovisual stimuli in randomized order, with each stimulus followed by three questions, in randomized order: 1) How much movement did you perceive? 2) How physically involved did you feel? and 3) How emotionally involved did you feel? Participants were asked to observe the stimuli and answer the questions as quickly and accurately as possible (no specific time limit was given), using the mouse to click on the Visual Analogue Scale (VAS) ranging from 0 (very little) to 100 (very much). The study design was the following: 2 Contours (Ascending, Still) * 3 Modalities (Audio, Video, Audiovisual), for a total of 6 conditions. Each experimental condition was repeated three times, with a total of 18 presented trials. Before carrying out the computer task, participants were presented with a brief training phase to become accustomed with the procedure. After the computer task, participants were asked to fill out a short debriefing survey about their experience. The experimental task was programmed using Psychopy 3.0 (Peirce, 2011) software.

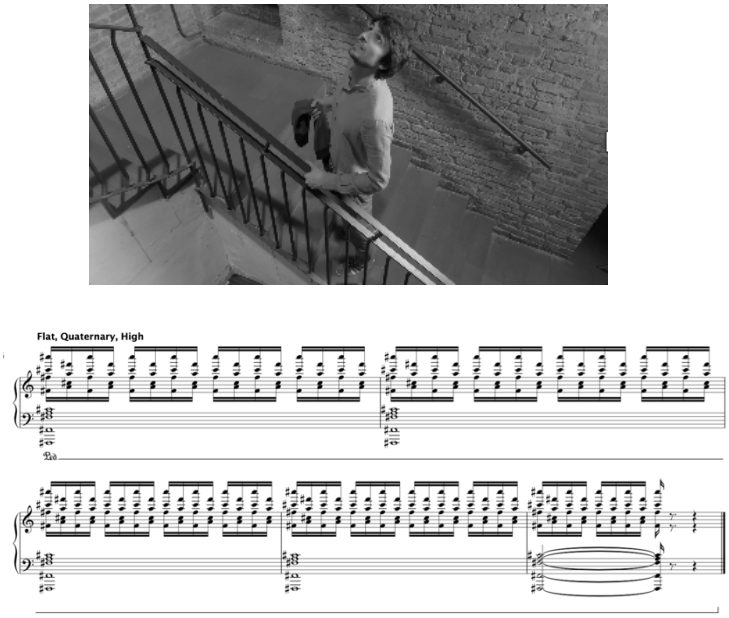
Stimulus	Contour	Modality	Complexity	Gender
1	Ascending	Audio	Low	n/a
2	Ascending	Audio	Medium	n/a
3	Ascending	Audio	High	n/a
4	Ascending	Video	n/a	Male
5	Ascending	Video	n/a	Male
6	Ascending	Video	n/a	Female
7	Ascending	Audiovisual	Low	Male
8	Ascending	Audiovisual	Medium	Male
9	Ascending	Audiovisual	High	Female
10	Still	Audio	Low	n/a
11	Still	Audio	Medium	n/a
12	Still	Audio	High	n/a
13	Still	Video	n/a	Female
14	Still	Video	n/a	Female
15	Still	Video	n/a	Male
16	Still	Audiovisual	Low	Female
17	Still	Audiovisual	Medium	Female
18	Still	Audiovisual	High	Male

Fig. 7.1 | Table of behavioral (audiovisual) study stimuli. Total number of stimuli: 27.



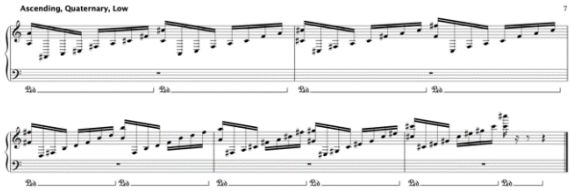
7.2

7.3

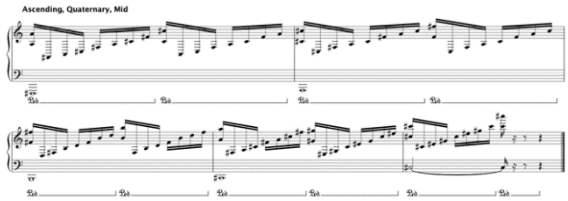


7.4

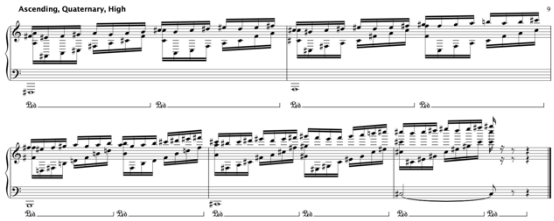
Fig. 7.2 – 7.4| Behavioral study: "Audiovisual Still" stimuli. "Still Female 1" with "Flat [Composition Contour], Quaternary [Note Pattern], Low [Composition Complexity]" (7.2); "Still Female 2" with "Flat, Quaternary, Middle" (7.3); "Still Male" with "Flat, Quaternary, High" (7.4).



7.5



7.6



7.7

Fig. 7.5 – 7.7 | Behavioral study: "Audiovisual Ascending" stimuli. "Ascending Male 1" with "Ascending [Composition Contour], Quaternary [Note Pattern], Low [Composition Complexity]" (7.5); "Ascending Male 2" with "Ascending, Quaternary, Middle" (7.6); and "Ascending Female" with "Ascending, Quaternary, High" (7.7).

7.2. EEG materials and methods

The aims of this experiment will be to investigate how unimodal and multimodal movement in the audiovisual image impact desynchronization and resynchronization (ERD/ERS) of the rolandic *mu* rhythm using high-density EEG and frequency analysis of 8 electrodes in each hemisphere around C3 and C4 sites. The hypotheses are the following: 1) [Ascending] conditions will have greater desynchronization of *mu* rhythm with respect to [Still] conditions, and 2) [Audiovisual] conditions will have greater desynchronization of *mu* rhythm with respect to [Video] and [Audio] conditions.

7.2.1. Stimuli

The same stimuli will be used as in the behavioral study. In total, 18 audiovisual monochrome stimuli will be used in each of the following modalities: Movement (Ascending, Still) and Modality (Audio, Video, Audiovisual).

7.2.2. Procedure

Participants will be seated in a shielded EEG lab in front of a computer screen placed on a table at a distance of 50 cm. The EEG will be recorded during 4 blocks, each separated by a short break. Each block will consist of 67-68 trials (in total 45 repetitions per condition, for a total number of 270 trials). Each trial will consist of a fixation cross of 500-1000ms (random length), followed by one of the 18 stimuli with a duration of 10,000ms. After stimulus presentation, a grey "resynchronization" screen will be presented for 3000-3500ms in order to enable signals to return to baseline. Participants will be asked to blink only in the second half of the resynchronization period in order to reduce movement artifacts. In 13.3% of trials, instead of the Audiovisual stimuli a red or blue ball will be presented, followed by a question asking whether the ball was red or blue, to which participants have to answer by clicking on the mouse positioned 15cm from the participants' right hand (action execution task). If the participant responds incorrectly or outside of the 2000ms time period, a message will appear on the screen stating that the response was either incorrect or given too late. There will be a total of 35 action execution/catch trials, counterbalanced and randomized for the color of the ball and the question (i.e., 25% with a blue/red combination, 25% with a red/blue combination, 25% with a blue/blue combination, and 25% with a red/red combination), giving

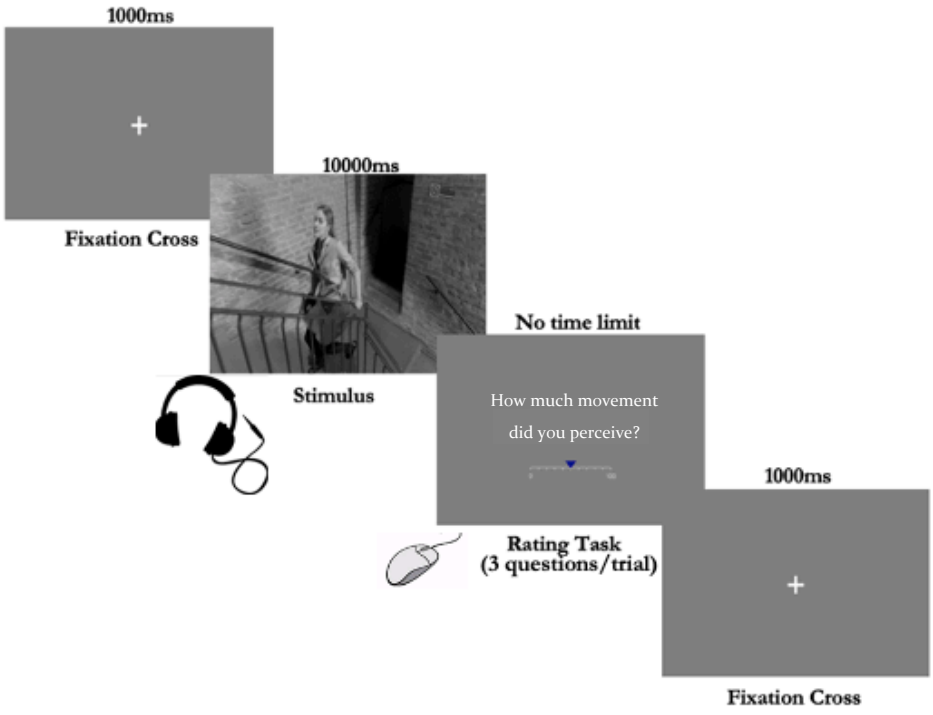


Fig. 7.8 | Example of behavioral study experimental trial. Components: fixation cross frame (1000ms), stimulus frame (10,000ms) and the Visual Analogue Scale (VAS) rating task (scale from 0 to 100, no time limit). Experiment was created using Psychopy 3.0.

a total of 35 catch trials, see Fig. 7.9. During the EEG recording, muscle responses from the hands will be measured using electromyography (EMG) in order to control for movements that may confound results in the observation trials. Stimuli will be presented using Psychopy 3 (Peirce, 2011). EEG data will be recorded using a 128-channel Sensor Net (Electrical Geodesic, Eugene, USA) using the standard EGI package Net Station 4.3.1. EEG raw data will be sampled at 500 Hz, electrode impedance will be set below 50 KΩ. The raw EEG data will be recorded using the vertex (Cz) as the online reference and will be re-referenced off-line with respect to the common average. EEG data processing will be conducted using the Matlab toolbox FieldTrip, and data will be filtered offline with a band-pass filter of 1-30 Hz, segmented into specific time epochs. An EMG signal will be recorded in all trials using an E.M.S. Triggerbox connected to a CED 1902 amplifier and Software for processing. In order to assess rolandic mu-rhythm modulation, time-frequency analysis will be performed on all EEG data for each participant on 1.5 second segments for all conditions. Based on previous studies, the following electrode clusters will be used: 8 electrodes in each hemisphere, localized around standard C3 and C4 sites (electrodes 30, 31, 36, 37, 41,42, 53, 54 on the left hemisphere and electrodes 79, 80, 86, 87, 93, 103, 104, 105 on the right hemisphere), see Fig. 7.10.

7.3. Behavioral results

In order to investigate whether VAS ratings are modulated by Contour and Modality, a linear mixed effect analysis was carried out. Participants' ratings were entered as dependent variables (Movement, Emotional Involvement and Physical Involvement), Contour (2 levels: Ascending, Still) and Modality (3 levels: Audio, Video, Audiovisual) as independent fixed variables, with participant intercepts as random effects. Tukey's test was used for post-hoc comparisons among means, where an error probability of less than 5% ($p < 0.05$) was considered significant in a population sample size of N=34. All analyses were performed using R software (R Core Team, 2019) and lme4 (Bates et al. 2015), ordinal (Christensen, 2019), effects (Fox, 2003) and emmeans (Lenth, 2020) functions; for data visualization, the ggplot2 package was used (Wickham, 2016).

7.3.1 Movement

The model explained 61.31% of the variance in Movement ratings, taking into account the random effects ($R^2_m = 0.45$, $R^2_c = 0.61$). The model revealed a significant main effect of Contour

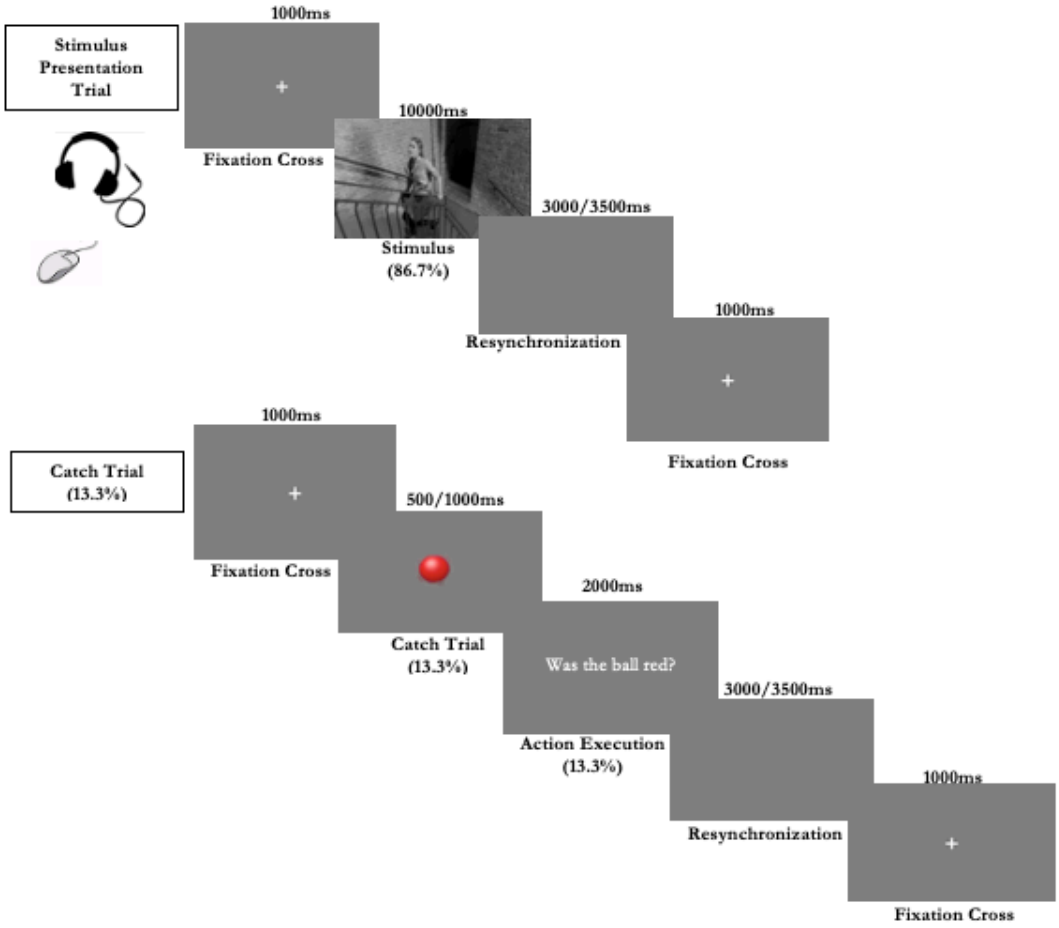


Fig. 7.9 | Example of EEG experimental trial. The "Stimulus Presentation Trial" constitutes 86.7% of total trials and includes the fixation cross frame (1000ms), the stimulus presentation (10,000ms) and the resynchronization frame (3000/3500ms); The "Catch Trial" constitutes 13.3% of total trials and include a fixation cross (1000ms), followed by a frame in which a red or blue ball is presented (50/1000ms), followed by the action execution task (i.e., answering the question with a mouse, time limit of 2000ms), and the resynchronization screen (3000/3500ms). Experiment was created using Psychopy 3.

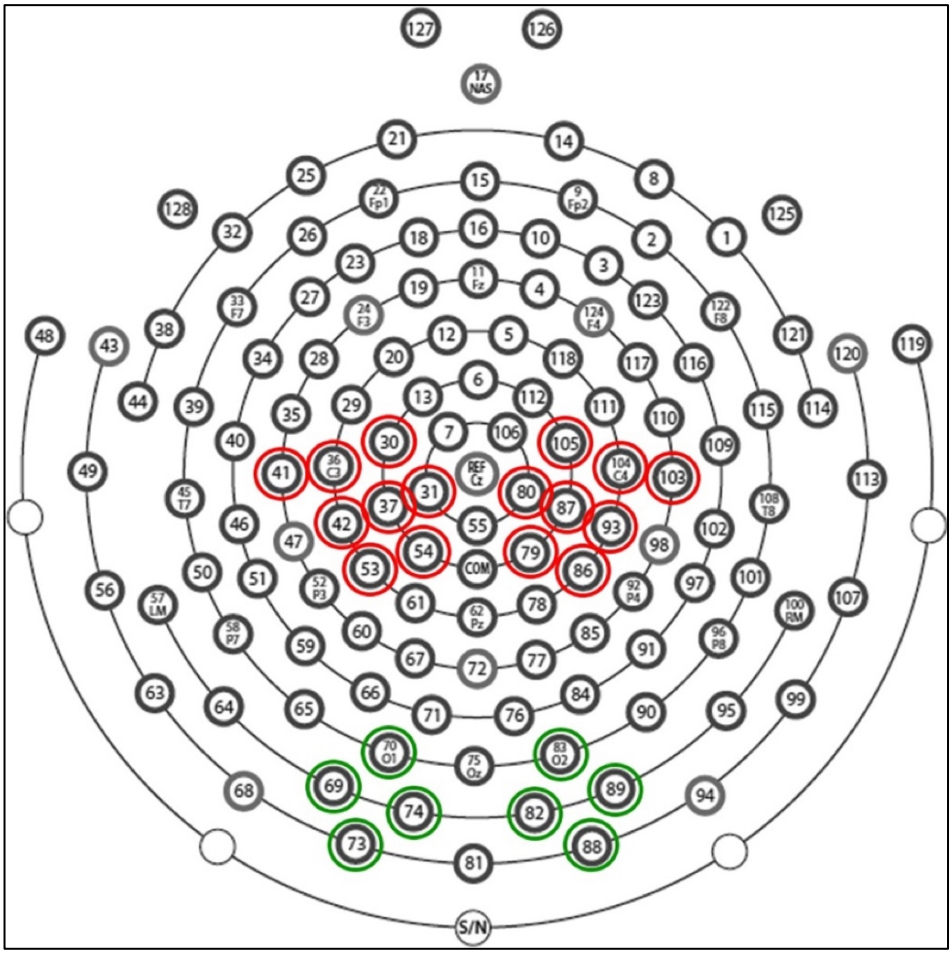


Fig. 7.10 | Hydrocel Geodesic Sensor Net– 128 channels. Electrodes circled in red depict central electrode-clusters chosen to investigate 136rolandic *mu*-rhythm modulation (i.e., electrodes 30, 31, 36, 37, 41,42, 53, 54 on the left and electrodes 79, 80, 86, 87, 93, 103, 104, 105 on the right). Electrodes circled in green indicate electro-clusters chosen to investigate whether recordings in central areas may be modulated by posterior activity (i.e., electrodes 69, 70, 73, 74 on the left and electrodes 82, 83, 88, 89 on the right). Source: Heimann et al., 2014.

($\chi^2_{(2)}$ =146.65, p < 0.0001), showing that participants perceived Ascending to have more Movement with respect to Still (z = 12.11, SE = 2.25; Moving: M = 71.8, SE =2.4; Still: M = 44.6, SE = 2.4). A significant main effect for Modality was found ($\chi^2_{(2)}$ = 16.32, p < 0.0001), showing that participants perceived Audio to have significantly more Movement than Video (z = 5.52, SE = 2.75; Audio: M = 64.6, SE = 2.65; Video: M = 49.4, SE = 2.65), and Audiovisual to have significantly more Movement with respect to Video (z = 4.03, SE = 2.75; Audiovisual: M = 60.5, SE = 2.65), with no difference between Audio and Audiovisual (z = 1.50, SE = 2.75). The model also showed a significant Contour*Modality interaction ($\chi^2_{(4)}$ = 28.21, p <0.0001), with post-hoc tests revealing that: Ascending Audio was perceived as having significantly less Movement than Ascending Audiovisual (z = -3.17, SE = 3.89; Ascending Audio: M = 66.4, SE = 3.29; Ascending Audiovisual: M = 78.7, SE = 3.29), no significant difference between Ascending Audiovisual and Ascending Video (z = 2.17, SE = 3.89; Ascending Video: M = 70.3, SE = 3.29), no significant difference between Ascending Audio and Ascending Video (z = -0.1, SE = 3.89), significantly more Movement in Still Audio than in Still Audiovisual (z = 5.28, SE = 3.89; Still Audio: M = 62.8, SE = 3.29; Still Audiovisual: M =42.3, SE = 3.29), significantly more Movement in Still Audiovisual than in Still Video (z = 3.53, SE = 3.89; Still Video: M =28.5, SE = 3.29), and significantly more Movement in Still Audio than in Still Video (z = 8.81, SE = 3.89). Post-hoc tests also revealed that: significantly more Movement was perceived in Ascending Audiovisual than in Still Audiovisual (z = 9.36, SE = 3.89), significantly more Movement in Ascending Video than in Still Video (z = 10.71, SE = 3.89), and no significant difference between Ascending Audio and Still Audio for perceived Movement (z = 0.91, SE = 3.89). See Fig. 7.11.

7.3.2 Physical Involvement

The model explained 58.19% of the variance in Physical Involvement ratings, taking into account the random effects (R^2_m = 0.12, R^2_c = 0.58). The model revealed a significant main effect of Contour ($\chi^2_{(2)}$ =25.00, p <0.0001), showing that participants perceived Ascending to be more Physically Involving with respect to Still (z = 5.0, SE = 2.28; Ascending: M = 48.0, SE = 3.33; Still: M = 36.6, SE = 3.33), see Figure 24. A significant main effect for Modality was found ($\chi^2_{(2)}$ = 14.30, p < 0.0001), showing that participants perceived Audio to be significantly more Physically Involving than Video (z = 4.55, SE = 2.79; Audio: M = 46.4, SE = 3.52; Video: M = 33.7, SE = 3.52), and Audiovisual to be significantly more Physically Involving than Video (z = 4.71, SE = 2.79; Audiovisual: M = 46.9, SE = 3.52), but no difference between Audio and Audiovisual (z = -0.16, SE = 2.79). The model also showed a significant Contour*Modality

interaction ($\chi^2_{(4)}$ = 3.18, p <0.05), with post-hoc tests revealing: no significant difference in Physical Involvement between Ascending Audio and Ascending Audiovisual (z = -1.77, SE = 3.94; Ascending Audio: M = 48.1, SE =4.04; Ascending Audiovisual: M = 55.1, SE =4.04), that Ascending Audiovisual was perceived as significantly more Physically Involving than Ascending Video (z = 3.58, SE =3.94; Ascending Video: M = 40.9, SE =4.04), no significant difference between Ascending Audio and Ascending Video (z = 1.81, SE = 3.94), no significant difference between Still Audio and Still Audiovisual (z = 1.54, SE = 3.94; Still Audio: M = 44.7, SE = 4.04; Still Audiovisual: M = 38.7, SE = 4.04), significantly more Physical Involvement in Still Audiovisual than in Still Video (z = 3.08, SE = 3.94; Still Video: M = 26.5, SE = 4.04), and significantly more Physical Involvement in Still Audio than in Still Video (z = 4.62, SE = 3.94). Post-hoc tests also revealed that: Ascending Audiovisual was significantly more Physically Involving than Still Audiovisual (z = 4.16, SE = 3.94), significantly more Physically Involving in Ascending Video than in Still Video (z = 3.65, SE = 3.94), and no significant difference between Ascending Audio and Still Audio (z = 0.85, SE = 3.94). See Fig. 7.11.

7.3.3 Emotional Involvement

The model explained 59.71% of the variance in Emotional Involvement ratings, taking into account the random effects (R^2_m = 0.20, R^2_c = 0.60). The model revealed a significant main effect of Contour ($\chi^2_{(2)}$ = 4.98, p < 0.05), showing that participants perceived Ascending to be more Physically Involving with respect to Still (z = 2.23, SE = 2.21; Ascending: M = 53.1, SE = 3.1; Still: M = 48.2, SE = 3.1). A significant main effect for Modality was found ($\chi^2_{(2)}$ = 48.21, p < 0.0001), showing that participants perceived Audio to be significantly more Emotionally Involving than Video (z = 9.39, SE = 2.71; Audio: M = 61.1, SE = 3.29; Video: M = 35.7, SE = 3.29), and Audiovisual to be significantly more Emotionally Involving than Video (z = 7.19, SE = 2.71; Audiovisual: M = 55.2, SE = 3.29), with no difference between Audio and Audiovisual (z = 2.20, SE = 2.71). See Fig. 7.11.

Contour and Modality

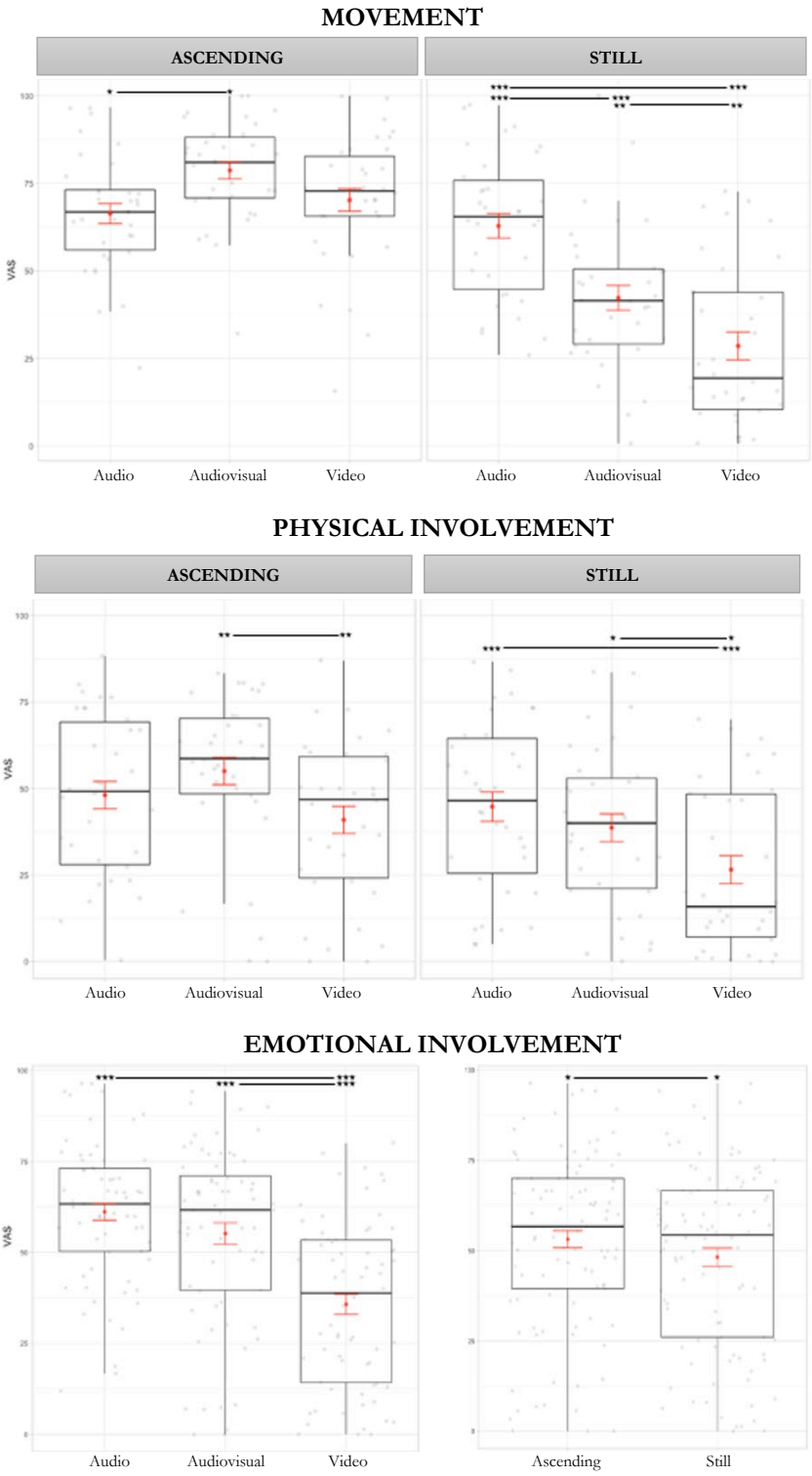


Fig. 7.11 | Boxplots for main effects: Contour and Modality. Boxplot depicting mean Visual Analogue Scale (VAS) ratings (on a scale from 0 to 100) for "Movement" with respect to main effects "Contour" and "Modality"; boxplot depicting mean VAS ratings (on a scale from 0 to 100) for "Physical Involvement" with respect to main effects "Contour" and "Modality; boxplot depicting mean VAS ratings (on a scale from 0 to 100) for "Emotional Involvement" with respect to main effects "Modality" and "Contour"; for all graphs, N=34 and error bars represent Standard Error (SE).

7.4. Discussion of behavioral study

Evidence in support of the theories of embodied simulation and embodied music cognition theory indicate that the audiovisual medium has the potential to evoke heightened forms of embodied experience (Gallese and Guerra, 2012, 2014, 2020; Cox, 2011, 2017). In the present study, we investigated the impact of Contour (Ascending, Still) and Modality (Audio, Video and Audiovisual) on participant ratings. The validated audio and video stimuli were selected to create 18 audiovisual stimuli for this experiment. The task was to rate the stimuli for perceived Movement and Emotional and Physical Involvement using a Visual Analogue Scale (VAS) ranging from 0 to 100, as quickly and as accurately as possible. As the interactions are our principal interest, the post-hoc test results will be discussed per dependent variable.

For Movement, results showed that Still Audio was perceived as having significantly more Movement than Still Audiovisual, Still Audiovisual was perceived as having significantly more Movement than Still Video, and Still Audio was perceived as having significantly more Movement than Still Video. It was also demonstrated that Ascending Audio was perceived as having significantly more Movement than Ascending Audiovisual, but no significant differences in Movement were found between Ascending Audiovisual and Ascending Video, and between Ascending Audio and Ascending Video. For Physical Involvement, results demonstrate that Still Audiovisual was perceived as being significantly more Physically Involving than Still Video, Still Audio was perceived as being significantly more Physically Involving than Still Video, but no significant difference was found between Still Audio and Still Audiovisual. It was also demonstrated that Ascending Audiovisual was perceived as being significantly more Physically Involving than Ascending Video, but not significant differences in Physical Involvement were found between Ascending Audio and Ascending Audiovisual, and between Ascending Audio and Ascending Video. For Emotional Involvement, only main effects were found. For the main effect Contour it was found that Ascending was perceived as being significantly more Emotionally Involving than Still. For the main effect Modality, it was found that Audio was perceived as being significantly more Emotionally Involving than Video, and Audiovisual as being significantly more Emotionally Involving than Video, but no significant differences were found between Audio and Audiovisual. Altogether, these results highlight the following. Firstly, it is demonstrated that Video and Audiovisual modalities are

modulated by Still vs. Ascending conditions for Movement and Physical Involvement ratings. Secondly, Audio Still is perceived as being significantly more Physically Involving than Video Still and as having significantly more Movement than Video Still and Audiovisual Still. Thirdly, in contrast to our hypothesis, Audio conditions are perceived as being significantly more Emotionally Involving than Video and Audiovisual conditions. Finally, Audio was not modulated by Still vs. Ascending conditions for Movement and Physical Involvement rating, contrary to our hypotheses.

Firstly, let us consider the results for Movement and Physical Involvement. Results demonstrate that whereas Contour (i.e., Ascending vs. Still) modulates Video and Audiovisual conditions, it does not modulate Audio conditions, contrary to our hypothesis. Similarly, Ascending Audiovisual and Ascending Video are perceived as significantly more Physically Involving than Still Audiovisual and Still Video, respectively, but no significant differences were found between Ascending Audio and Still Audio. A possible explanation for this is that Still Audio conditions were not really perceived as "still". That is, in Still Video, the actor is not moving side to side – the image is "still" in both "horizontal" and "vertical" planes. However, in Still Audio, the notes oscillate between two intervals, and this note density may be perceived as "kinetic impressions" by participants, i.e., as movement. The reason that Still Audio tracks are not ideally "still", a limitation of the study, is because it was necessary to control for note density. If the Still Audio tracks were made to be a resounding, single note, which on a piano is difficult to achieve due to its rapid decay, Ascending and Descending audio tracks would have had a far greater note density in comparison, and therefore it would not have been possible to compare Movement ratings between Ascending, Descending and Still tracks in the Music Validation. The high ratings attributed to perceived Movement for Audio Still reverberate in the finding that overlaying Still Video with Still Audio to create Still Audiovisual conditions significantly reduces ratings for perceived Movement. In Ascending conditions, on the other hand, overlaying Ascending Video with Ascending Audio to create Ascending Audiovisual significantly increases ratings for perceived Movement, whereas no significant differences are reported between Ascending Audiovisual and Ascending Video. For Physical Involvement, however, the opposite was true – Ascending Audiovisual is perceived as significantly more Physically involving than Ascending Video, but not more than Ascending Audio. This may be explained again by the primacy, or bias, for visual motion being

more “explicit” than melodic motion. What is also interesting to note is that, regardless, Still Audio and Ascending Audio conditions were given higher ratings for Movement (Still Audio: M = 62.8; Ascending Audio: M = 66.4), whereas Ascending Video ratings had a mean of 70.3. This is significant as it suggests that when Audio is unimodal, it is perceived as being highly embodied, but when it is presented multimodally with Video, on an explicit level, visual motion drives the attention of the beholder.

With respect to Emotional Involvement, no interactions were found between Contour and Modality, contrary to our hypothesis. The main effect result that Audiovisual is perceived as significantly more Emotionally Involving than Video, but not than Audio, may be explained by the low valence of the actors’ facial expressions which drove participants’ ratings. A limitation of the study, it was necessary to direct the actors to keep their expressions “expressively indistinct” (i.e., emotionally ambiguous), to borrow Johannes Riis’ term, in order to control for distinct facial expressiveness influencing appraisal of the videos, rather than the presence of movement/contour in the image. In his examination of the pre-Tsarist star Ivan Mozzhukhin’s acting style and its relevance to the Kuleshov effect, Riis considers the emotional range of (in)expressive (in)distinctness (Riis, 2018; Kolesnikov, 2020h). The effect of an indistinct, emotionally ambiguous facial expression is effective when associated to a “rich” and “distinct” context, which in turn resolves the ambiguity. In the absence of such an explicit context in the video stimuli of this study, which begin and end in the course of an action without a clear objective, the emotionally ambiguous expressions of the actors are unresolved, and thus emotionally uninvolving. This element stands as one of the principal differences between our stimuli and the original scene from *Zhuravli* (1957), in which Boris (Batalov) displayed a distinctively expressive excitedness and joy as he dashed up the stairs toward Veronika (Samoilova).

As a whole, these results provide further evidence in support of, 1) the notion that embodied cognition, vitality forms and kinetic qualities are as relevant to musical experience as they are to visual experience, and 2) the role of (in)distinct expressiveness of acting and (in)distinctness of contextual cues in driving emotional involvement with a scene. Further studies, in particular implicit studies, are needed to investigate the impact of contour movement on sensorimotor activation and emotional activation.

8. Conclusion

In this thesis, a neurohumanistic approach has been adopted in order to investigate a "constellation" of the cultural notion of gesture, both visual, musical and audiovisual, as explored in the exegesis of Eisenstein's vast body of work as well as in the scholarship of other historical and contemporary thinkers, and how these notions of "audio/visual" gesture have manifested both in the cinema of Eisenstein and beyond it. Eisenstein's initial triangulation of gesture has been "re-triangulated" and operationalized in an in-depth investigation of the cognitive and neural underpinnings of audio/visual engagement, mediated through deep interdisciplinary collaboration which, until now, has been lacking.

In Chapter 1, the notion of the "actor's gesture" is explored. It is illustrated how Meyerhold's principles of biomechanics, an amalgamation of notions from Delsarte's and Dalcroze's ideas on eurhythmics and James and Bekhterev's principles of objective psychology, in turn were formalized by Eisenstein into an integrated conceptual framework on expressive movement and subsequently applied to film theory and practice, as embodied in his early treaty, "Montazh Attraksionov" (1924). It is also examined how Eisenstein's ideas on expressive movement relate to the "kinaesthetic turn" at the end of the nineteenth century, when the advent of cinema and "visual culture" emerged out of the "loss of gesture" during the 19th century. Such views are examined in the scholarship of Freeburg, Phillips, Sargent, Balázs, Benjamin, and Brecht as well as in the contemporary work of Geil, Baumbach and Agamben.

In Chapter 2, the notion of gesture is explored within the context of musical multimedia such as opera and cinema. Specifically, "musical gesture" is examined in relation to Eisenstein's theories of vertical montage and embodiment of emotional theme in "Vertikalny'i Montazh" (1940), in particular, formulated in the aftermath of his collaboration with Prokofiev on *Nevskii* (1938) and in the foreground of his staging of *Die Walküre* in 1940. Wagner's polyphonous gesture and *Gesamtkunstwerk* are explored in the context of his anti-capitalism and the 19th century crisis of lost gesture, and Eisenstein's account of Wagner's legacy in his essay "Voploshenie mifa" (1940) are examined, informed by Puchner's and Bartlett's analyses. Alongside the discussion of gesture as the link between sound and image, Adorno's, Horkheimer's and Eisler's critique of the culture industry and Eisenstein's appropriation of Wagner's ideas on the "total work of art" and polyphonous gesture is examined.

In Chapter 3, the notion of "gesture as audiovisual link" is expanded in relation to Eisenstein's writings on organic unity, synesthesia, regression, pathos, and the *Gründproblem* of art in *Metod* (2002), *Neravnodushnaia priroda* (1940) and "Opredeliaushii zhest" (1940). In particular, informed by Schulzki's analysis, d'Udine's influence on Eisenstein's ideas about dance and rhythm as the precursors to music are examined, and the "oprichnina" scene in *Ivan Chast' II* (1958) as the apotheosis of Eisenstein's ideas on gesture and synesthesia, informed by Khitrova's analysis, is discussed. The gesture's expansion into multimodality in *Ivan* in particular in the semiotic and metric modes of "turning", is contrasted with examples in other films, specifically in *The Red Shoes* (1948), *Moi laskovyi i nezhnyi zver'* (1978) and *The Shawshank Redemption* (1994).

In Chapter 4, it is reported how, in order to facilitate the triangulation of film theory and cognitive and neural sciences in the investigation of visual, musical and audiovisual gestures in film experience, a novel contribution to methodology was implemented through two artistic collaborations. As an aesthetic model for the construction of naturalistic, custom-made and carefully controlled stimuli, the staircase scene from the Soviet film *Zhuravli* (1957) was used for its dynamic embodiment of the motif of "rising" (achieved through the upward turning of the camera and actor bodies). The scene was recreated using a DJI Phantom 4 Drone with Fabrizio LaPalombara, and original musical tracks were recorded by the film composer Eduardo Andrade.

The theoretical considerations in Chapters 1 to 3 elucidate Eisenstein's overlap with the contemporary frameworks of embodied simulation and embodied music cognition. This intercross is operationalized in Chapters 5 to 7 through the use of behavioral methods. Due to the ongoing Covid-19 pandemic, the neurophysiological investigation using EEG had to be postponed.

In Chapter 5, a Validation of the video stimuli is reported, consisting of the following aims: 1) to investigate the effect of Camera Movement (Ascending, Descending, Still), Actor Presence (Female, Male, None [No Actor]) and Image Speed (Normal, Slow, Very Slow) on participant ratings, and 2) to validate and select stimuli with the highest ratings for subsequent behavioral and EEG experiments. The reported results on Actor Presence, in which the presence of human movement (Female and Male conditions) evoke a greater sense of Physical

Involvement and Movement with respect to the absence of human movement (No actor conditions), provide additional support for embodied simulation theory, which predicts that we resonate more with conspecifics and familiar motor repertoires. Although the absence of "distractions" may increase "immersion" in the scene, resulting in identification with the drone and simulation of its ascending rotation in space, participants found the presence of an actor in the video image to be more involving on a motor level. Further neuroimaging studies are needed to shed light on this result. The reported results on Camera Movement also provide support for embodied simulation theory, and are in line with prior neuroimaging studies on motor cortex activation during the observation of human and camera movement.

The reported distinction between the perception of Ascending (i.e., rising) gestures and Descending and Still gestures is of particular interest. In addition to providing further evidence for embodied simulation theory, it provides support for Eisenstein's notion of "embodiment of emotional theme" and the primacy of the "underlying gesture" as an expressive means in the cinema and the arts in general. As Eisenstein writes, "the physical movement of a man gives rise to a particular excitation for that movement. Thus, laughter evokes a good mood, tears - sadness. William James" (Eisenstein, 1964: 136). Eisenstein extended the gesture in the actor to shot composition as a whole in his theory of vertical montage where "[…] the human being and the interaction between his *gesture* and *intonation* when they are generated by a single emotion, will prove to be the determining prototype of shaping audiovisual structures […]" (Eisenstein, 2010: 328). From rising gestures in the "pictorial composition of the shot", Kalatozov, an ex-student of Eisenstein's at VGIK, used the camera as a space for gestural compositions. Indeed, this result provides "empirical" validation for Kalatozov's creative decision to express Boris' last happy memory before the war through the semiotic mode of "upness", and the vitality form and embodied metaphor of "joy as rising". The motif of "rising" is embodied not only in Boris' body, but through the rotating ascent of the camera, which we simulate. Indeed, participants rate Ascending videos as more Emotionally Involving than Descending and Still (i.e., static) videos, and they rate them as evoking a greater sensation of Movement as well, indicative of a motor engagement in addition to the recognition of the semiotic mode of "upness" and "rising". A possible explanation may be the increased perception of effort, or exertion, associated with ascending the stairs, as opposed to descending movements, despite both movements being goal-

orientated. That there is a parallel distinction in the reported results of the Music Validation is also of particular interest.

In Chapter 6, a validation of the music stimuli is reported, consisting of the following objectives: 1) to investigate the effect of Composition Contour (Ascending, Descending, Still), Composition Complexity (Low, Medium and High) and Note Pattern (Binary, Ternary and Quaternary) on participant ratings, and 2) to validate and select stimuli with the highest ratings for subsequent behavioral and EEG experiments. The reported results on Composition Contour demonstrate not only that the presence of Composition Contour correlates with higher participant ratings, but that Ascending melodies, like Ascending videos, are perceived as evoking more Movement and Emotional Involvement than Descending or Flat Melodies; furthermore, they are perceived as being more Physically Involving than Ascending or Flat Melodies. This result provides support for the mimetic subvocalization theory of Cox's embodied music cognition framework, specifically the Mimetic Motor Imagery and Voice as Source Domain hypotheses, in which greater perceived exertion of singing higher notes, for example, is theorized to result in greater mimetic participation and thus stronger musical affect. The results of the two Validations indicate a cross-modal mapping for "ascending – exertion – affect" across visual and musical gestures. In addition to Composition Contour, the results reported for Composition Complexity provide further support for the notion that musical notes are as perceived as kinetic events, and that higher note density in a melody correlates with more gestures needed to create the relevant kinetic impressions. This is visible in the results for Note Pattern where, in contrast to predictions, Quaternary melodies are perceived as evoking significantly more Rotation, Movement and Physical Involvement than Ternary melodies, calling for further studies to shed light on the relation between note density and perceived Rotation. Notably, it was found that Ascending melodies are also perceived as evoking more Rotation in participants with respect to Descending and Flat melodies. This result is particularly interesting in light of the rising and rotating gesture of the camera in the "staircase scene", which future studies on multimodal gesture could investigate further.

The results of the Music Validation and their grounding in Cox's theory of embodied music cognition strongly relate to Eisenstein's notion of voice as the "forerunner" of melody and the "source" from which we draw the "basic human emotional content" in a musical melody

(Eisenstein, 2020: 228). In his incorporation of Wagner's notion of "leitmotif" and d'Udine's notion of the "universal gesture" pervading all arts, Eisenstein's model of musical experience is decidedly kinesthetic, in which it is possible to "visualize certain graphic images" and depict "with a movement of our hand, the impression a certain detail of a piece of music makes on us" (Eisenstein, 2010: 374). In his films *Nevskii* and *Ivan*, Eisenstein infuses the multimodal space of his films with the "kinetic impressions" of Prokofiev's score, in which the musical gesture inhabits forms of vitality on par with those of the actor.

In Chapter 7, following the Validation and selection of the musical and video stimuli, the aims of the behavioral (audiovisual) study were to investigate how Contour (Ascending, Still) and unimodal vs. multimodal Modality (Audio, Video and Audiovisual) influenced participant ratings (the investigation of sensorimotor resonance using electroencephalography (EEG) had to be postponed). The reported results demonstrate that, firstly, Video and Audiovisual modalities are modulated by Still as opposed to Ascending conditions for Movement and Physical Involvement ratings. Secondly, Audio Still is perceived as being significantly more Physically Involving than Video Still and as having significantly more Movement than Video Still and Audiovisual Still. Thirdly, contrary to our predictions, Audio conditions are perceived as being significantly more Emotionally Involving than Video and Audiovisual conditions. Finally, Audio conditions are not modulated by Still as opposed to Ascending conditions for Movement and Physical Involvement ratings, contrary to our hypotheses. With respect to Emotional Involvement, the main effect result that Audiovisual is perceived as significantly more Emotionally Involving than Video, but not than more Emotionally Involving than Audio, may be explained by the low valence of the actors' facial expressions which drove participants' ratings. A limitation of the study, it was necessary to direct the actors to keep their expressions "expressively indistinct" (i.e., emotionally ambiguous) in order to control for effects on participants' appraisal of features related to motion. The matter of distinct facial expressiveness is an important difference between our stimuli and the original staircase scene from *Zhuravli,* in which Batalov's (i.e., Boris) acting conveys explicit excitement and joy as ran up the stairs to meet Samoilova (Veronika). As a whole, the results of the behavioral study provide further evidence in support of the notion that embodied gesture and kinetic qualities are as relevant to our perception of music as they are to our perception of visual scenes, as elaborated in Eisenstein body of work.

There are several avenues for future research in the third-culture space explored in this thesis. Subsequent to the completion of the EEG study, a key piece of the puzzle, a more in-depth investigation of the effect of multimodal correspondence between gestures could be pursued. Specifically, an investigation of how formal features in single modalities influence the degree of "hybridization", to use Levenson's term, between sound and image could be conducted. For example, as preliminarily addressed in Chapter 3, why do certain rotating or rising multimodal gestures feel "smoother" than others? Behavioral and neuroimaging studies examining how variation in melodic contour and note pattern influence the perception of "synthesis", or fusion, when combining music with camera movement, could provide key insight on the links between embodied perception and multimodality in the experience of film.

In conclusion, the contents of this thesis reflect the initial stages of a neurohumanistic, third culture approach to the notion of gesture and its sensorimotor basis in audiovisual experience. There are still numerous lines of research to be undertaken, as well as challenges underlying interdisciplinary research to be overcome. However, the hard interdisciplinary work of collaboration with experts in the arts, humanities, and cognitive science will serve to bridge the gap between the mind, brain, and mediated representations such as audiovisual experience. In the pursuit of a third-cultural space for the study of cinema, it is unlikely that Eisenstein will cease to be a source of inspiration in the aesthetic project of neurohumanism.

9. Acknowledgements

My sincere gratitude first of all to my research supervisors, Prof. Michele Guerra and Prof. Vittorio Gallese, for giving me the opportunity to pursue this research, and for their invaluable guidance and support throughout the course of this project; to Prof. Maria Alessandra Umiltà, for her continuous mentorship, instruction and moral support in carrying out empirical research; and to Fabrizio La Palombara and Eduardo Andrade for providing their invaluable knowledge and artistic insight during our creative collaborations. My thanks also to my colleagues at Prof Gallese's research lab, as well as to Francesca Siri, Gioacchino Garofalo, Katalin Balint, Tim Smith, Joshua Bamford, Joe Dixon, Katrin Heimann and Joerg Fingerhut for their crucial support and insights. My thanks to my friends and to Michele Serrazanetti, for his love and support. And last, but certainly not least, I am profoundly grateful to my sister and parents, in particular, to whom I owe everything.

10. Bibliography

Abreu, A. M., Macaluso, E., Azevedo, R. T., Cesari, P., Urgesi, C. and Aglioti, S. M. (2012) 'Action anticipation beyond the action observation network: a functional magnetic resonance imaging study in expert basketball players', *European Journal of Neuroscience*. Wiley Online Library, 35(10), pp. 1646–1654.

Adorno, T. (2005) *In Search of Wagner*, Trans. Rodney Livingstone. With a Foreword by Slavoj Žižek. London & New York: Verso.

Adorno, T. W. and Eisler, H. (2007) *Composing for the Films*. Later Edition. London: Continuum.

Agamben, G. (1993) *Infancy and History: The Destruction of Experience*. Verso.

Agamben, G. (2007) *Profanations*. Translated by J. Fort; Brooklyn, NY: Zone Books.

Aglioti, S. M., Cesari, P., Romani, M. and Urgesi, C. (2008) 'Action anticipation and motor resonance in elite basketball players', *Nature Neuroscience*, 11(9).

Albrecht, H. and Wöllner, C. (2015) 'Metaphors and Embodied Meaning in Film Music: The Case of Leitmotifs', in Fahlenbrach, K. (ed.) *Embodied Metaphors in Film, Television, and Video Games: Cognitive Approaches*. 1st edition. New York ; London: Routledge, pp. 145–162.

Allen, R. W. (1987) 'The Aesthetic Experience of Modernity: Benjamin, Adorno, and Contemporary Film Theory', *New German Critique*, (40), pp. 225-240.

Ardizzi, M., Calbi, M., Tavaglione, S., Umiltà, M. A. and Gallese, V. (2020) 'Audience spontaneous entrainment during the collective enjoyment of live performances: physiological and behavioral measurements', *Scientific Reports*. Nature Publishing Group, 10(1), p. 3813. doi: 10.1038/s41598-020-60832-7.

Aristotle, *Aristotle - Problemata, Loeb Classical Library*. Available at: https://www.loebclassics.com/view/aristotle-problems/2011/pb_LCL317.275.xml (Accessed: 18 December 2020).

Auerbach, J. (2007) *Body Shots: Early cinema's incarnations.* University of California Press.

Aumont, J. and Hildreth, L. (1983) 'Montage Eisenstein I: Eisensteinian Concepts', *Discourse*. Wayne State University Press, 5, pp. 41–99.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G. and Iacoboni, M. (2006) 'Congruent Embodied Representations for Visually Presented Actions and Linguistic Phrases Describing Actions', *Current Biology*, 16(18), pp. 1818–1823. doi: 10.1016/j.cub.2006.07.060.

Babiloni, C., Babiloni, F., Carducci, F., Cincotti, F., Cocozza, G., Del Percio, C., Moretti, D. V. and Rossini, P. M. (2002) 'Human Cortical Electroencephalography (EEG) Rhythms during

the Observation of Simple Aimless Movements: A High-Resolution EEG Study', *NeuroImage*, 17(2), pp. 559–572. doi: 10.1006/nimg.2002.1192.

Balázs, B. (2010) *Early Film Theory : Visible Man and The Spirit of Film*. Berghahn Books.

Balázs, B. (2012) *Theory of the Film*. Ulan Press.

Bartlett, R. (1992) 'The Embodiment of Myth: Eizenshtein's Production of "Die Walküre"', *The Slavonic and East European Review*, pp. 53-76.

Bateman, J.A. and Schmidt, K.H. (2013) *Multimodal film analysis: how films mean*. New York: Routledge (Routledge studies in multimodality, 5).

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2014) 'Fitting linear mixed-effects models using lme4', *arXiv preprint arXiv:1406.5823*.

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K. and Jancke, L. (2007) 'A network for audio–motor coordination in skilled pianists and non-musicians', *Brain Research*, 1161, pp. 65–78. doi: 10.1016/j.brainres.2007.05.045.

Baumbach, N. (2019) *Cinema/politics/philosophy*. Columbia University Press.

Bekhterev, V. M. (1913) *La psychologie objective* [Objective Psychology]. Librairie F.Alcan.

Belodubrovskaya, M. (2018) 'Eisenstein's Attractions, Mirror Neurons, and Contemporary Action Cinema', *Projections*, 12(1), pp. 1-18.

Benjamin, W. (1998) 'What is Epic Theatre?', in *Understanding Brecht*. London; New York: Suhrkamp Verlag, pp. 1–15.

Benjamin, W. (1999) 'Franz Kafka - On the Tenth Anniversary of His Death', in *Selected Writings - Volume 2 - 1927-1934*. Belknap Press of Harvard University Press.

Benjamin, W. (2005) 'Surrealism: The Last Snapshot of the European Intellegentsia', in Jennings, M., Eiland, H., and Smith, G. (eds) *Walter Benjamin: Selected Writings, Volume 2: Part 1*. Belknap Press of Harvard University Press.

Benjamin, W. (2008) 'The Work of Art in the Age of Its Technological Reproducibility - Second Version', in Jennings, M., Doherty, B., and Levin, T. (eds) *The Work of Art in the Age of Its Technological Reproducibility*. Harvard University Press, pp. 19–56. doi: 10.7312/kul-17602-004.

Biesen, S. C. (2015) '"Kinesthesis" and Cinematic Montage: An Historical Examination of the Film Theories and Avant-Garde Mediation of Slavko Vorkapich in Hollywood', *Studies in Visual Arts and Communication: an International Journal*, 2(1), pp. 1–10.

Bode, R. (1922) *Ausdrucksgymnastik* [Expressive Gymnastics]. Beck.

Bordwell, D. (2005) *The Cinema of Eisenstein*. 1st edition. New York, NY: Routledge.

Bordwell, D. (1980) 'The Musical Analogy', *Yale French Studies*, (60), pp. 141-156

Brown, S. W. (1995) 'Time, change, and motion: The effects of stimulus movement on temporal perception', *Perception & psychophysics*. Springer, 57, pp. 105–116.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G. and Freund, H.-J. (2001) 'Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study', *European Journal of Neuroscience*, 13(2), pp. 400–404. doi: https://doi.org/10.1111/j.1460-9568.2001.01385.x.

Bulgakowa, O. (2002) *Sergei Eisenstein. a Biography*. Berlin: PotemkinPress.

Bulgakowa, O. (2014) 'From expressive movement to the "basic problem". The Vygotsky-Luria-Eisensteinian theory of Art', *The Cambridge handbook of cultural-historical psychology*, pp. 423–448.

Bull, M. and Back, L. (2020) *The Auditory Culture Reader*. Routledge.

Calbi, M., Aldouby, H., Gersht, O., Langiulli, N., Gallese, V. and Umiltà, M. A. (2019) 'Haptic Aesthetics and Bodily Properties of Ori Gersht's Digital Art: A Behavioral and Eye-Tracking Study', *Frontiers in Psychology*, 10, p. 2520. doi: 10.3389/fpsyg.2019.02520.

Callan, D. E., Tsytsarev, V., Hanakawa, T., Callan, A. M., Katsuhara, M., Fukuyama, H. and Turner, R. (2006) 'Song and speech: brain regions involved with perception and covert production', *Neuroimage*. Elsevier, 31(3), pp. 1327–1342.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E. and Haggard, P. (2005) 'Action Observation and Acquired Motor Skills: An fMRI Study with Expert Dancers', *Cerebral Cortex*. Oxford Academic, 15(8), pp. 1243–1249. doi: 10.1093/cercor/bhi007.

Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E. and Haggard, P. (2006) 'Seeing or Doing? Influence of Visual and Motor Familiarity in Action Observation', *Current Biology*, 16(19), pp. 1905–1910. doi: 10.1016/j.cub.2006.07.065.

Carluccio, G. A. and Villa, F. (2006) *Il corpo del film. Scritture, contesti, stile, emozioni* [The Body of Film. Writings, contexts, styles, emotions]. Carocci Editore.

Carocci, E. (2018) *Il sistema schermo-mente* [The Screen-Mind System]. Bulzoni, Roma.

Cervini, A. (2008) *Sulla biomeccanica. Azione scenica e movimento* [On Biomechanics. Stage Action and Movement]. Roma: Armando Editore.

Chen, J. L., Penhune, V. B. and Zatorre, R. J. (2008a) 'Listening to musical rhythms recruits motor regions of the brain', *Cerebral cortex*. Oxford University Press, 18(12), pp. 2844–2854.

Chen, J. L., Penhune, V. B. and Zatorre, R. J. (2008b) 'Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training', *Journal of cognitive neuroscience*. MIT Press, 20(2), pp. 226–239.

Chion, M., Gorbman, C. and Murch, W. (1994) *Audio-vision: sound on screen*. New York: Columbia University Press.

Christensen, R. H. B. (2019) 'A Tutorial on fitting Cumulative Link Mixed Models with clmm2 from the ordinal Package', *Tutorial for the R Package ordinal https://cran. r-project. org/web/packages/ordinal/Accessed*, 1.

Christie, I. and Scorsese, M. (2002) *Arrows of Desire*. London ; Boston: Ian Christie with Martin Scorsese.

Claude Lelouch à propos de 'Quand passent les cigognes' [Claude Lelouch on 'The Cranes Are Flying'] (2015). Available at: https://www.youtube.com/watch?v=vG-DfdOy7SI&t=320s (Accessed: 2 December 2020).

Coëgnarts, M. and Kravanja, P. (eds) (2015) *Embodied Cognition and Cinema*. Leuven: Leuven University Press.

Cook, N. (1998) *Analysing Musical Multimedia*. Oxford: Oxford University Press.

Core, T. R. (2013) *R: A language and environment for statistical computing*. Vienna, Austria.

Cox, A. (2001) 'The mimetic hypothesis and embodied musical meaning', *Musicae scientiae*. SAGE Publications Sage UK: London, England, 5(2), pp. 195–212.

Cox, A. (2011) 'Embodying music: Principles of the mimetic hypothesis', *Music Theory Online*, 17(2).

Cox, A. (2017) *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking*. Illustrated edition. Bloomington; Indianapolis: Indiana University Press.

Crowther, B. (1960) *Screen: Exchange Film: Soviet 'Cranes Are Flying' Bows Here - The New York Times*, *The New York Times*. Available at: https://www.nytimes.com/1960/03/22/archives/screen-exchange-filmsoviet-cranes-are-flying-bows-here.html (Accessed: 2 December 2020).

D'Aloia, A. and Eugeni, R. (2014) 'Neurofilmology. An Introduction', *Cinema & Cie*. Available at: https://iris.unicampania.it/handle/11591/400317#.X8fPEaozbOQ (Accessed: 2 December 2020).

Das Rheingold | work by Wagner. *Encyclopedia Britannica*. Available at: https://www.britannica.com/topic/Das-Rheingold (Accessed: 12 December 2020).

Davis, M. H. (1980) 'A multidimensional approach to individual differences in empathy'. American Psychological Association Washington, DC.

Davison, A. (2004) *Hollywood theory, non-Hollywood practice: cinema soundtracks in the 1980s and 1990s*. Aldershot, Hants, England ; Burlington, VT: Ashgate (Ashgate popular and folk music series).

Deathridge, J. (1983) 'Review of In Search of Wagner', *19th-Century Music*. University of California Press, 7(1), pp. 81–85. doi: 10.2307/746549.

Der Ring Des Nibelungen Leitmotifs [The Ring of Nibelung Leitmotifs] (2014) Available at: https://www.youtube.com/playlist?list=PL78TsyiiZjhGNl-civwjVsk_7tn6XG3wh (Accessed: 19 December 2020).

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. and Rizzolatti, G. (1992) 'Understanding motor events: a neurophysiological study | SpringerLink', *Experimental Brain Research*, (91), pp. 176–180.

Diamond, D. (2016) 'The Red Shoes: A fairy tale within a ballet within a film.', *Psychoanalytic Psychology*, 33(Suppl 1), pp. S104–S119. doi: 10.1037/pap0000071.

Didi-Huberman, G. (2020) 'La danza di ogni cosa' (The dance of all things), in A. Somaini (ed)*Time machine. Vedere e sperimentare il tempo* (Time machine. Seeing and experimenting with time). *Ediz. illustrata*. Milano: Skira, pp. 83–102.

Doga, E. (1978/2020) *Waltz*. https://eugendoga.musicaneo.com/.

Doga, E. (2015) *Waltz*. Available at: https://www.youtube.com/watch?v=vM7P4AnHILI (Accessed: 16 December 2020).

Dujardin, É. (1887) *Revue wagnérienne* [The Wagnerian Review], *Gallica*. Available at: https://gallica.bnf.fr/ark:/12148/bpt6k199971 (Accessed: 24 December 2020).

Darabont, F. (1994) *The Shawshank Redemption.* United States: Columbia Pictures.

Easdale, B. (1948) *The Red Shoes. Composed and Arranged for Piano Solo by Brian Easdale.* United Kingdom: Chappell.

Eduardo Andrade Composer. Available at: https://www.eduardoandrademusic.com (Accessed: 30 December 2020).

Eisenstein, S. (1925) *Bronenosets potemkin* [The Battleship Potemkin]. USSR: Mosfilm.

Eisenstein, S. (1938) *Aleksandr Nevskii* [Alexander Nevsky]. USSR: Mosfilm.

Eisenstein, S. (1940) 'Pered prem'eroi "Val'kirii"'[Before the premiere of "The Valkyrie"], *Vecherniaia Moskva*, 21 September 1940, p. 3.

Eisenstein, S. (1958) *Ivan Groznyi Chast' II* [Ivan the Terrible Part II]. USSR: Mosfilm.

Eisenstein, S. (1964) *Izbronnye proizvedeniia* [Selected Works]. Edited by S. Yutkevich. Moscow: Iskusstvo.

Eisenstein, S. (1968-74) *Izbronnye proizvedeniia* [Selected Works]. Edited by S. Yutkevich. Moscow: Iskusstvo.

Eisenstein, S. (1988) *Nonindifferent Nature: Film and the Structure of Things*. First Edition. Translated by H. Marshall. Cambridge; New York: Cambridge University Press.
Eisenstein, S. (1991) *Kinovedcheskie Zapiski* [Film Studies Notes], (11/41).

Eisenstein, S. (2002) *Metod* [Method]. Edited by N. Kleiman. Moscow: Eisenstein-Tsentr, Muzei Kino.

Eisenstein, S. (2004) *Neravnodushnaia Priroda Tom 1* [Nonindifferent Nature Volume 1]. Moscow: Eisenstein Center.

Eisenstein, S. (2005) *Eizenshtein o Meierholde, 1919-1948* [Eisenstein on Meyerhold, 1919-1948]. Moscow: Novoye Izdatel'stvo.

Eisenstein, S. (2006) *The Eisenstein Collection*. Seagull Books.

Eisenstein, S. (2010a) *Writings, 1922-1934: Sergei Eisenstein Selected Works, Volume 1*. London ; New York: I.B. Tauris.

Eisenstein, S. (2010b) *Towards a Theory of Montage: Sergei Eisenstein Selected Works, Volume 2*. Annotated edition. Edited by M. Glenny and R. Taylor. London ; New York: I.B. Tauris.

Eisenstein, S. (2010c) *Writings, 1934-1947: Sergei Eisenstein Selected Works, Volume 3*. Reprint edition. Edited by R. Taylor. Translated by W. Powell. London ; New York: I.B. Tauris.

Eisenstein, S. (2014) *Immoral Memories: An Autobiography*. Peter Owen Publishers.

Eitan, Z. (1997) *Highpoints: A study of melodic peaks*. University of Pennsylvania Press.

Elcott, N.M. (2016) *Artificial Darkness: An Obscure History of Modern Art and Media*. University of Chicago Press.

Engell, L. and Siegert, B. (eds) (2016) *Schwerpunkt Verschwinden*. Hamburg: Felix Meiner Verlag (Zeitschrift für Medien- und Kulturforschung, Heft 7, 1 (2016)).

Epstein, J. (1947) *Le cinéma du diable* [The Cinema of the Devil]. J. Melot.

Eugeni, R. (2018) *Semiotica dei media. Le forme dell'esperienza* [Media semiotics. Forms of experience]. 6th edn. Roma: Carocci.

Fabrizio La Palombara. Available at: https://www.fabriziolapalombara.com (Accessed: 30 December 2020).

Faul, F., Erdfelder, E., Lang, A.-G. and Buchner, A. (2007) 'G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences', *Behavior research methods*. Springer, 39(2), pp. 175–191.

Fox, J. (2003) 'Effect displays in R for generalised linear models', *Journal of statistical software*, 8(15), pp. 1–27.

Freeburg, V.O. (1918) *The Art of Photoplay Making*. The Macmillan company. Available at: http://archive.org/details/artphotoplaymako1freegoog (Accessed: 12 December 2020).

Freedberg, D. and Gallese, V. (2007) 'Motion, emotion and empathy in esthetic experience', *Trends in Cognitive Sciences*, 11(5), pp. 197–203. doi: 10.1016/j.tics.2007.02.003.

Frith, U. and Frith, C. (2010) 'The social brain: allowing humans to boldly go where no other species has been', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, 365(1537), pp. 165–176. doi: 10.1098/rstb.2009.0160.

Gallese, V., Fadiga, L., Fogassi, L. and Rizzolatti, G. (1996) 'Action recognition in the premotor cortex', *Brain*. Oxford Academic, 119(2), pp. 593–609. doi: 10.1093/brain/119.2.593.

Gallese, V. and Goldman, A. (1998) 'Mirror neurons and the simulation theory of mind-reading', *Trends in Cognitive Sciences*, 2(12), pp. 493–501. doi: 10.1016/S1364-6613(98)01262-5.

Gallese, V. (2001) 'The shared manifold'hypothesis. From mirror neurons to empathy', *Journal of consciousness studies*. Imprint Academic, 8(5–6), pp. 33–50.

Gallese, V. (2005) 'Embodied Simulation: From Neurons to Phenomenal Experience.' *Phenomenology and the Cognitive Sciences* 4: 23-48.

Gallese, V. (2007) 'Before and Below Theory of Mind: Embodied Simulation and the Neural Correlates of Social Cognition'. *Philosophical Transactions of the Royal Society of London B* 362: 659-669.

Gallese, V. (2009) 'Motor Abstraction: a Neuroscientific Account of How Action Goals and Intentions Are Mapped and Understood.' *Psychol Res* 73(4): 486-498.

Gallese, V. (2014) 'Bodily Selves in Relation: Embodied Simulation as Second-Person Perspective on Intersubjectivity.' *Phil. Trans. R. Soc. B* 69, 20130177, published 28 April 2014.

Gallese, V. and Guerra, M. (2014) 'The feeling of motion: Camera movements and motor cognition', *Cinema & Cie*, 14(22–23), pp. 103–12.

Gallese, V. and Guerra, M. (2012) 'Embodying movies: Embodied simulation and film studies'. *Cinema: Journal of Philosophy and the Moving Image*, *3*, pp.183-210.

Gallese, V. and Guerra, M. (2020) *The Empathic Screen: Cinema and Neuroscience*, *The Empathic Screen*. Oxford University Press.

Gallese, V., Keysers, C. and Rizzolatti, G. (2004) 'A unifying view of the basis of social cognition', *Trends in Cognitive Sciences*, 8(9), pp. 396–403. doi: 10.1016/j.tics.2004.07.002.

Gallese, V. and Sinigaglia, C. (2011) 'What is so special about embodied simulation?', *Trends in Cognitive Sciences*, 15(11), pp. 512–519. doi: 10.1016/j.tics.2011.09.003.

Gallez, D. W. (1978) 'The Prokofiev-Eisenstein Collaboration: "Nevsky" and "Ivan" Revisited', *Cinema Journal*, 17(2), p. 13. doi: 10.2307/1225488.

Gastaut, H. J. and Bert, J. (1954) 'EEG changes during cinematographic presentation (Moving picture activation of the EEG)', *Electroencephalography and Clinical Neurophysiology*, 6, pp. 433–444. doi: 10.1016/0013-4694(54)90058-9.

Gaudreault, A. and Gunning, T. (2006) 'Early cinema as a challenge to film history', *The cinema of attractions reloaded*. JSTOR, pp. 365–380.

Geil, A. (2018) 'Between gesture and physiognomy: "universal language" and the metaphysics of film form in Béla Balázs' *Visible Man*', *Screen*, 59(4), pp. 512–522. doi: 10.1093/screen/hjy054.

Gibbon, M. (1948) *The red shoes ballet: A critical study*. 2nd edition. Saturn Press.

Gibson, J. J. (1966) 'The senses considered as perceptual systems.' Houghton Mifflin.

Gibson, J. J. (1979) *The ecological approach to visual perception. Boston, MA, US*. Houghton, Mifflin and Company.

Global Cinematography Institute. *Renowned DPs Who Never Won an Oscar*. Available at: https://www.globalcinematography.com/gci-journal/2020/2/13/renwoned-dps-who-never-won-an-oscar (Accessed: 30 December 2020).

Godøy, R. I. and Leman, M. (eds) (2010) *Musical Gestures: Sound, Movement, and Meaning*. 1st edition. New York: Routledge.

Goswami, U. (2015) 'Sensory theories of developmental dyslexia: three challenges for research', *Nature Reviews Neuroscience*. Nature Publishing Group, 16(1), pp. 43–54.

Grespi, B. (2019) *Figure del corpo. Gesto e immagine in movimento* [Figures of the body. Gesture and image in movement]. Milano: Meltemi.

Grodal, T. (2009) *Embodied Visions: Evolution, Emotion, Culture, and Film*. Oxford University Press.

Gunning, T. (1986) 'The Cinema of Attractions: Early Film, Its Spectator and the Avant-Garde', *Wide Angle*, 8(3–4), pp. 63–70.

Halpern, A. R. and Zatorre, R. J. (1999) 'When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies', *Cerebral cortex*. Oxford University Press, 9(7), pp. 697–704.

Halpern, A. R., Zatorre, R. J., Bouffard, M. and Johnson, J. A. (2004) 'Behavioral and neural correlates of perceived and imagined musical timbre', *Neuropsychologia*. Elsevier, 42(9), pp. 1281–1292.

Hansen, M.B. (2011) *Cinema and Experience: Siegfried Kracauer, Walter Benjamin, and Theodor W. Adorno*. First Edition. University of California Press.

Hansen, M. B. (1981) 'Introduction to Adorno, "Transparencies on Film" (1966)', *New German Critique*, (24/25), p. 186. doi: 10.2307/488049.

Hansen, M. B. (2004) 'Room-for-Play: Benjamin's Gamble with Cinema', *October*. The MIT Press, 109, pp. 3–45.

Hansen, N. C. and Huron, D. (2019) 'Twirling triplets: the qualia of rotation and musical rhythm', *Music & Science*. SAGE Publications Sage UK: London, England, 2, p. 2059204318812243.

Hanslick, E. (1891) *The beautiful in music: a contribution to the revisal of musical aesthetics*. Novello.

Hari, R. and Kujala, M. V. (2009) 'Brain Basis of Human Social Interaction: From Concepts to Brain Imaging', *Physiological Reviews*. American Physiological Society, 89(2), pp. 453–479. doi: 10.1152/physrev.00041.2007.

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N. and Heeger, D. J. (2008) 'Neurocinematics: The neuroscience of film', *Projections*. Berghahn Journals, 2(1), pp. 1–26.

Heimann, K., Uithol, S., Calbi, M., Umiltà, M. A., Guerra, M., Fingerhut, J. and Gallese, V. (2019) 'Embodying the camera: An EEG study on the effect of camera movements on film spectators′ sensorimotor cortex activation', *PLOS ONE*. Edited by A. Avenanti, 14(3), p. e0211026. doi: 10.1371/journal.pone.0211026.

Heimann, K., Umiltà, M. A. and Gallese, V. (2013) 'How the motor-cortex distinguishes among letters, unknown symbols and scribbles. A high density EEG study', *Neuropsychologia*, 51(13), pp. 2833–2840. doi: 10.1016/j.neuropsychologia.2013.07.014.

Heimann, K., Umiltà, M. A., Guerra, M. and Gallese, V. (2014) 'Moving Mirrors: A High-density EEG Study Investigating the Effect of Camera Movements on Motor Cortex Activation during Action Observation', *Journal of Cognitive Neuroscience*. MIT Press, 26(9), pp. 2087–2101. doi: 10.1162/jocn_a_00602.

Heimann, K. S., Uithol, S., Calbi, M., Umiltà, M. A., Guerra, M. and Gallese, V. (2017) '"Cuts in Action": A High-Density EEG Study Investigating the Neural Correlates of Different Editing Techniques in Film', *Cognitive science*. Wiley Online Library, 41(6), pp. 1555–1588.

Hickok, G., Buchsbaum, B., Humphries, C. and Muftuler, T. (2003) 'Auditory–motor interaction revealed by fMRI: speech, music, and working memory in area Spt', *Journal of cognitive neuroscience*. MIT Press, 15(5), pp. 673–682.

Horkheimer, M. and Adorno, T. W. (2002) *Dialectic of Enlightenment*. 1st edition. Edited by G. S. Noerr. Translated by E. Jephcott. Stanford, Calif: Stanford University Press.

Hufner, M. (1998) '"Composing for the Films" (1947): Adorno, Eisler and the sociology of music', *Historical Journal of Film, Radio and Television*, 18(4), pp. 535–540. doi: 10.1080/01439689800260341.

Iablonska, M. *O biomehanike Meierhol'da* (2014) Available at: http://ziernie-performa.net/blog/2014/01/22/malgozhata-yablonska-o-biomexanike-mejerxolda/ (Accessed: 12 December 2020).

Iampolski, M. (1991) 'Kuleshov's Experiments and the New Anthropology of the Actor.', in Taylor, R. and Christie, I. (eds) *Inside the Film Factory: New Approaches to Russian and Soviet Cinema*.

Iampolski, M. (2016) 'Point - Pathos - Totality', in Kleiman, N. I., Somaini, A., and Eisenstein, S. (eds) *Sergei M. Eisenstein: notes for a general history of cinema*. Amsterdam: Amsterdam University Press (Film theory in media history), pp. 357–373.

Ivanov, V. V. (1976) 'Essays on the History of Semiotics in the USSR', *Moscow: Languages of Russian Culture*.

Ivan the Terrible, Op. 116: Pt. 2 The Tsar's Banquet & The Cathedral, Song Of The Oprichniki I - YouTube (2016) Available at: https://www.youtube.com/watch?v=M_u1-Aszs8I&list=RDiODOrAzJjus&index=3 (Accessed: 2 December 2020).

James, W. (1884) 'What is an Emotion?', *Mind*. Oxford University Press, Mind Association, 9(34), pp. 188–205.

Johnson, M. (2007) *The Meaning of the Body: Aesthetics of Human Understanding*. Reprint edition. Chicago London: University of Chicago Press.

Kalatozov, M. (1957) *Letiat Zhuravli* [The Cranes Are Flying]. USSR: Mosfilm.

Khitrova, D. (2011) 'Eisenstein's choreography in Ivan the Terrible', *Studies in Russian and Soviet Cinema*, 5(1), pp. 55–71. doi: 10.1386/srsc.5.1.55_1.

Kleiman, N. (1993) 'Arguments and Ancestors', in Christie, I. and Taylor, R. (eds) *Eisenstein Rediscovered*. 1st edn. Routledge, pp. 32–40.

Koelsch, S. (2012) *Brain and Music*. John Wiley & Sons.

Kohler, E. (2002) 'Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons', *Science*, 297(5582), pp. 846–848. doi: 10.1126/science.1070311.

Kolesnikov, A. (2020a) *Stimulus 1*. Available at: https://www.youtube.com/watch?v=PaUWr4Bw2jc (Accessed: 31 December 2020).

Kolesnikov, A. (2020b) *Stimulus 2*. Available at: https://www.youtube.com/watch?v=hIIijlXrClM (Accessed: 31 December 2020).

Kolesnikov, A. (2020c) *Stimulus 3*. Available at: https://www.youtube.com/watch?v=-Tx7VoV_7zE (Accessed: 31 December 2020).

Kolesnikov, A. (2020d) *Stimulus 4*. Available at: https://www.youtube.com/watch?v=SuHzVN3Zs5Y (Accessed: 31 December 2020).

Kolesnikov, A. (2020e) *Stimulus 5*. Available at: https://www.youtube.com/watch?v=jCeUEgFaG28 (Accessed: 31 December 2020).

Kolesnikov, A. (2020f) *Stimulus 6*. Available at: https://www.youtube.com/watch?v=eqg7LBrMIbY (Accessed: 31 December 2020).

Kolesnikov, A. (2020g) *The Cranes Are Flying (Kalatozov, 1957) - The Staircase Scene*. Available at: https://www.youtube.com/watch?v=TRcZbexAoTU (Accessed: 3 December 2020).

Kolesnikov, A. (2020h) 'The Geocultural Provenance of Narratives: The Case of the Kuleshov Effect', *Film History*. Indiana University Press, 32(2), pp. 55–79. doi: 10.2979/filmhistory.32.2.03.

Konigsberg, I. (2007) 'Film Studies and the New Science', *Projections*, 1(1), pp. 1–24. doi: 10.3167/proj.2007.010102.

Kracauer, S. (1927) 'Bücher vom Film [Books from Film]', *Frankfurter Zeitung*.

Kracauer, S. (1995) *The Mass Ornament: Weimar Essays*. Harvard University Press.

Krausshouse (2011) *Wagner's Ring Cycle Leitmotifs, Metropolitan Opera Orchestra*. Available at: https://www.youtube.com/watch?v=wvA54D2Gz3E (Accessed: 18 December 2020).

Kristensen, S. B., Sandberg, K. and Bibby, B. M. (2020) 'Regression methods for metacognitive sensitivity', *Journal of Mathematical Psychology*. Elsevier, 94.

Lakoff, G. and Johnson, M. (2008) *Metaphors We Live By*. University of Chicago Press.

Langer, S. K. (1957) *Problems of Art*. Charles Scribner's Sons.

Law, A. and Gordon, M. (eds) (1996) *Meyerhold, Eisenstein, and Biomechanics: Actor Training in Revolutionary Russia*. Mcfarland.

Lenth, R., Singmann, H., Love, J., Buerkner, P. and Herve, M. (2020) *Emmeans: Estimated marginal means, aka least-squares means (1.4. 5)[Computer software]*.

Letiat Zhuravli [The Cranes Are Flying] (2016) Available at: https://vid1.ria.ru/ig/infografika/01/project01/shadoof_bird/index.html (Accessed: 2 December 2020).

Levaco, R. (1973) 'The Eisenstein-Prokofiev Correspondence', *Cinema Journal*. [University of Texas Press, Society for Cinema & Media Studies], 13(1), pp. 1–16. doi: 10.2307/1225055.

Levinson, J. (2011) *Music, art, and metaphysics*. Oxford University Press.

Lifton, M. (1960) 'Review: The Cranes Are Flying by Mikhail Kalatozov', *Film Quarterly*. University of California Press, 13(3), pp. 42–44. doi: 10.2307/1210437.

Lipps, T. (1906) *Ästhetik: psychologie des schönen und der kunst* [Aesthetics: psychology of beauty and art]. Voss.

Loteanu, E. (1978) *Moi laskovyi i nezhnyi zver'* [The Hunting Accident]. USSR: Mosfilm.

MacKay, J. (2007) 'Film Energy: Process and Metanarrative in Dziga Vertov's The Eleventh Year (1928)', *October*. MIT Press, pp. 41–78.

Mallarmé, S. (1976) *Un Coup de dés* [A Roll of the Dice]. Available at: http://www.gallimard.fr/Catalogue/GALLIMARD/Poesie-Gallimard/Igitur-Divagations-Un-Coup-de-des (Accessed: 24 December 2020).

McLean, A. L. (2008) *Dying Swans and Madmen: Ballet, the Body, and Narrative Cinema*. New Brunswick, N.J: Rutgers University Press.

Meyerhold, V. (1914) 'Tehnika tsennitcheskovo dvijénie [Technique of Stage Movement]', *L'amour de trois oranges*, (4–5).

Meyerhold, V. (1921) 'Ob uiazvimah mestah v teatral'novo fronta [On vulnerable spots in the theatrical front]', *Vestnik Teatra*, (72–73).

Michotte, A. (2017) *The Perception of Causality*. Routledge.

MIT Libraries. *Motion capture of pole vaulting | 150 Years in the Stacks*. (2011) Available at: https://libraries.mit.edu/150books/2011/02/07/1892/1892-ill2/ (Accessed: 28 December 2020).

Moisei Vainberg – Pis'ma o Tashkente [Moisey Vaynberg – Letters on Tashkent]. Available at: https://mytashkent.uz/2016/09/17/moisej-vajnberg/ (Accessed: 30 December 2020).

Musescore. *Sergei Prokofiev: Music for Sergei Eisenstein's Film Ivan the Terrible, op. 116: Song of the Oprichniki*, *Musescore.com*. Available at: https://musescore.comhttps://musescore.com/user/11930751/scores/4990289 (Accessed: 28 December 2020).

Muthukumaraswamy, S. D., Johnson, B. W. and McNair, N. A. (2004) 'Mu rhythm modulation during observation of an object-directed grasp', *Cognitive Brain Research*, 19(2), pp. 195–201. doi: 10.1016/j.cogbrainres.2003.12.001.

Münsterberg, H. (2014) *The Photoplay: A Psychological Study*. The Floating Press.

Ndebele, P. (2013) 'The Declaration of Helsinki, 50 years later', *Jama*. American Medical Association, 310(20), pp. 2145–2146.

Neuberger, J. (2003) *Ivan the Terrible: The Film Companion*. Bloomsbury Academic.

Neuberger, J. (2017) 'Another Dialectic: Eisenstein on Acting', in Neuberger, J., Somaini, A., and Kleĭman, N. I. (eds) *The flying carpet: studies on Eisenstein and Russian cinema in honor of Naum Kleiman*. London: Thames & Hudson, pp. 255-279.

Neuberger, J. (2019) *This Thing of Darkness: Eisenstein's Ivan the Terrible in Stalin's Russia*. Illustrated edition. Ithaca New York: Cornell University Press.

Nicholls, M. (2018) 'A Group Marriage without Sex: Fusion and Collaboration in The Archers' The Red Shoes (1948)', *Journal of Film and Video*. [University of Illinois Press, University Film & Video Association], 70(1), pp. 32–46. doi: 10.5406/jfilmvideo.70.1.0032.

Nietzsche, F. (2003) *The Birth of Tragedy: Out of the Spirit of Music*. Translated by Shaun Whiteside. Penguin UK, eBook (https://play.google.com/store/books/details/Friedrich_Nietzsche_The_Birth_of_Tragedy?id=JJL6Xsr1ztwC).

Nietzsche, F. W. (2005) *Thus spoke Zarathustra: A book for everyone and nobody*. Oxford University Press, USA.

Noyes, R. and Kletti, R. (1976) 'Depersonalization in the Face of Life-Threatening Danger: A Description', *Psychiatry*. Washington, etc., United States: William Alanson White Psychiatric Foundation., 39(1), pp. 19–27.

Oldfield, R. C. (1971) 'The assessment and analysis of handedness: the Edinburgh inventory', *Neuropsychologia*. Oxford, 9(1), pp. 97–113.

Olenina, A. H. (2020) *Psychomotor Aesthetics: Movement and Affect in Modern Literature and Film*. Oxford, New York: Oxford University Press.

Olenina, A. H. and Schulzki, I. (2017) 'Mediating Gesture in Theory and Practice', *Apparatus. Film, Media and Digital Cultures of Central and Eastern Europe*, 0(5). doi: 10.17892/app.2017.0005.100.

O'Pray, M. (1993) 'The frame and montage in Eisenstein's "later" aesthetics', in Chrisite, I. and Taylor, R. (eds) *Eisenstein Rediscovered*. 1st edn. Routledge, pp. 32–40.

Orgs, G., Dombrowski, J.-H., Heil, M. and Jansen-Osmann, P. (2008) 'Expertise in dance modulates alpha/beta event-related desynchronization during action observation', *European Journal of Neuroscience*, 27(12), pp. 3380–3384. doi: https://doi.org/10.1111/j.1460-9568.2008.06271.x.

Overy, K. (2003) 'Dyslexia and music: From timing deficits to musical intervention', *Annals of the New York Academy of Sciences*. Blackwell Publishing Ltd Oxford, UK, 999(1), pp. 497–505.

Patel, A. D. (2009) 'Snowball the Dancing Cockatoo'. IUPUI (Campus). Office of Communications and Marketing.

Peatman, M. M. (1976) 'Sergei Eisenstein's "Ivan the Terrible" As a Cinematic Realization of the Concept of the "Gesamtkunstwerk"'. PhD Dissertation.

Peirce, J., Gray, J., Halchenko, Y., Britton, D., Rokem, A. and Strangman, G. (2011) 'PsychoPy–A Psychology Software in Python'.

Phillips, H. A. (1914) *The Photodrama*. New York, Arno Press.

Pinotti, A. (2016) *L'empathie. Histoire d'une idée de Platon au posthumain* [Empathy. The history of an idea from Plato to posthuman]. Paris: VRIN.

Pinotti, A. and Somaini, A. (2012) *Aura e choc - Saggi sulla teoria dei media* [Aura and shock - Essays on media theory]. Einaudi.

Pinotti, A. and Somaini, A. (2016) *Cultura visuale. Immagini, sguardi, media, dispositivi* [Visual culture. Images, gazes, media, devices] Vol. 653. Einaudi.

Plato. *Plato, Republic, Book 2*. Available at: http://www.perseus.tufts.edu/hopper/text?doc=Plat.+Rep.+2 (Accessed: 3 December 2020).

Powell, M. and Pressburger, E. (1948) *The Red Shoes*. United Kingdom: The Archers.

Powell, M. (2000) *A Life in Movies by Michael Powell*. Faber and Faber.
Pratt, C. C. (1968) *The Meaning of Music - A Study in Psychological Aesthetics*. Johnson Reprint Corporation.

Prokofiev - Alexander Nevsky 5 The Battle on Ice. (2011) Available at: https://www.youtube.com/watch?v=-0RbStmxvm4 (Accessed: 18 December 2020).

Puchner, M. (1999) 'Polyphonous Gestures: Wagnerian Modernism from Mallarmé to Stravinsky', *Criticism*. Wayne State University Press, 41(1), pp. 25–39.

Rebecchi, M., Vogman, E., and Gathmann, T. (2017) *Sergei Eisenstein and the Anthropology of Rhythm*. Nero.

Repp, B. (1992) 'Music as Motion: A Synopsis of Alexander Truslit's (1938)', *Haskins Laboratories Status Report on Speech Research*, pp. 265–278.

Riis, J. (2018) 'Ivan Mozzhukhin's Acting Style: Beyond the Kuleshov Effect', *Projections*. Berghahn Journals, 12(2), pp. 137–158.

Rilke, R. M. (1969) *Letters, 1910-1926*. Norton.

Rizzolatti, G., Fadiga, L., Gallese, V. and Fogassi, L. (1996) 'Premotor cortex and the recognition of motor actions', *Cognitive Brain Research*, pp. 131–141.

Roberts, R., Callow, N., Hardy, L., Markland, D. and Bringer, J. (2008) 'Movement imagery ability: development and assessment of a revised version of the vividness of movement

imagery questionnaire', *Journal of Sport and Exercise Psychology*. Human Kinetics, Inc., 30(2), pp. 200–221.

Robertson, R. (2009) *Eisenstein on the Audiovisual: The Montage of Music, Image and Sound in Cinema*. I.B.Tauris. doi: 10.5040/9780755697168.

Rogers, S. (2013) 'Truth, Lies, and Meaning in Slow Motion Images', in Shimamura, A. P. *Psychocinematics: Exploring Cognition at the Movies*. Illustrated edition. Oxford, New York: Oxford University Press, pp. 149-164.

Rosen, P. (2020) 'Adorno and Film Music: Theoretical Notes on Composing for the Films', p. 27.

Rouse, A. A., Cook, P. F., Large, E. W. and Reichmuth, C. (2016) 'Beat keeping in a sea lion as coupled oscillation: implications for comparative understanding of human rhythm', *Frontiers in neuroscience*. Frontiers, 10, p. 257.

Ruprecht, L. (2015) 'Gesture, Interruption, Vibration: Rethinking Early Twentieth-Century Gestural Theory and Practice in Walter Benjamin, Rudolf von Laban, and Mary Wigman', *Dance Research Journal*. Cambridge University Press, 47(2), pp. 23–41. doi: 10.1017/S0149767715000200.

Sargent, E. W. (1916) *Technique of the Photoplay*. New York, The Moving Picture World. Available at: http://archive.org/details/techniquephotopoosarggoog (Accessed: 12 December 2020).

Sbriscia-Fioretti, B., Berchio, C., Freedberg, D., Gallese, V. and Umiltà, M. A. (2013) 'ERP Modulation during Observation of Abstract Paintings by Franz Kline', *PLOS ONE*. Public Library of Science, 8(10), p. e75241. doi: 10.1371/journal.pone.0075241.

Schenker, H. (1954) *Harmony by Heinrich Schenker*. University of Chicago Press.

Schreibman, M. A. and Steiner, M. (2004) 'On Gone with the Wind, Selznick, and the Art of "Mickey Mousing": An Interview with Max Steiner', *Journal of Film and Video,* 56(1), pp. 41-50.

Schulzki, I. (2019) 'Chapter Seven "The Underlying Gesture": Towards the Notion of Gesture in Jean d'Udine | Semantic Scholar', in Catanese, R., Lavina, F.S. and V. Valente (eds) *From Sensation to Synaesthesia in Film and New Media*. Cambridge Scholars Publishing.

Sessions, R. (1941) *The Composer and His Message In: The Intent of the Artist. Edited by A. Centeno*. Princeton: Princeton University Press.

Score's Animation (2016) *Sull' aria - Le Nozze di Figaro* [The Wedding of Figaro] - *W. A. Mozart Score Animation*. Available at: https://www.youtube.com/watch?v=aUf6zTjfb9s (Accessed: 18 December 2020).

Shaw, G. B. (2010) *The Perfect Wagnerite*. 4th ed. edition. New York: Dover Publications.

Shimamura, A. P. (2013) *Psychocinematics: Exploring Cognition at the Movies*. Illustrated edition. Oxford, New York: Oxford University Press.

Shpolberg, M. (2017) 'Humanising the Soviet subject: *letiat zhuravli (The Cranes Are Flying)*', *Sesnses of Cinema* (85), p. 7.

Siri, F., Ferroni, F., Ardizzi, M., Kolesnikov, A., Beccaria, M., Rocci, B., Christov-Bakargiev, C. and Gallese, V. (2018) 'Behavioral and autonomic responses to real and digital reproductions of works of art', *Progress in Brain Research*. Edited by J. F. Christensen and A. Gomila. (The Arts and The Brain), 237, pp. 201–221. doi: 10.1016/bs.pbr.2018.03.020.

Siri, F., Kolesnikov, A., Ardizzi, M., Calbi, M., Ferroni, F., Ferro, A. M., Freedberg, D., Umiltà, M. A. and Gallese, V. (2020) 'The Perceptual Experience of Portraits and Self-Portraits: A Behavioural and Eye-Tracking Study', *SSRN Electronic Journal*. doi: 10.2139/ssrn.3519045.

Sirotkina, I. (2014) *Lektsiia "Fizicheskii teatr avangarda: biomehanika, ekstsentrika, slovoplastika" Irina Sirotkina* [Lecture "Avant-garde physical theatre: biomechanics, eccentricity, wordplay" | Irina Sirotkina], Evreiskii muzei i Tsentr tolerantnosti [Jewish Museum and Centre for Tolerance], Moscow. https://www.youtube.com/watch?v=U90PMQI6ScE

Sirotkina, I. (2019) 'The science and art of movement: French and Russian contributions to the studies of human action", *Preprint*, https://www.academia.edu/24235649/The_art_and_science_of_movement_in_France_and_Russia.

Smart, M. A. (2004) *Mimomania: music and gesture in nineteenth-century opera*. Berkeley: Univ. of California Press.

Smith, M. (2017) *Film, art, and the third culture: a naturalized aesthetics of film*. First edition. Oxford ; New York, NY: Oxford University Press.

Smith, T. (2014) 'Audiovisual correspondences in Eisenstein's Alexander Nevsky: a case study of viewer attention', in Taberham, P. (ed.) *Cognitive Media Theory*. Routledge. doi: 10.4324/9780203098226.

Snow, C. P. (1998) *The Two Cultures*. Cambridge University Press.

Sobchack, V. (1982) 'Toward inhabited space: The semiotic structure of camera movement in the cinema', *Semiotica*. De Gruyter Mouton, 41(1–4), pp. 317–335. doi: 10.1515/semi.1982.41.1-4.317.

Sobchack, V. (2006) '"Cutting to the Quick": Techne, Physis, and Poiesis and the Attractions of Slow Motion', in Strauven, W. (ed.) *The Cinema of Attractions Reloaded*. Amsterdam University Press, pp. 337–352.

Somaini, A. (2016) 'Cinema as "Dynamic Mummification," History as Montage: Eisenstein's Media Archaeology', in Kleiman, N. I., Somaini, A., and Eisenstein, S. (eds) *Sergei M.*

Eisenstein: notes for a general history of cinema. Amsterdam: Amsterdam University Press (Film theory in media history), pp. 19–109.

Sonkusare, S., Breakspear, M. and Guo, C. (2019) 'Naturalistic Stimuli in Neuroscience: Critically Acclaimed', *Trends in Cognitive Sciences*, 23(8), pp. 699–714. doi: 10.1016/j.tics.2019.05.004.

Sparks, R., Helm, N. and Albert, M. (1974) 'Aphasia rehabilitation resulting from melodic intonation therapy', *Cortex*. Elsevier, 10(4), pp. 303–316.

Stadler, J. (2018) 'Between Movies and Mind, Affective Neuroscience, and the Philosophy of Film', *Projections,* 12(2), pp. 86-94.

Stafford, B. M. (2011) *A Field Guide to a New Meta-field: Bridging the Humanities-Neurosciences Divide*. University of Chicago Press.

Stern, D. N. (2010) *Forms of Vitality: Exploring Dynamic Experience in Psychology and the Arts*. Illustrated edition. Oxford ; New York: Oxford University Press.

Streltsova, A., Berchio, C., Gallese, V. and Umiltà, M. A. (2010) 'Time course and specificity of sensory-motor alpha modulation during the observation of hand motor acts and gestures: a high density EEG study', *Exp Brain Res*, p. 11.

Tarasti, E. (1978) *Myth and Music: A Semiotic Approach to the Aesthetics of Myth in Music, Especially that of Wagner, Sibelius and Stravinsky*. Suomen Musiikkitieteellinen Seura.

The Metropolitan Museum of Art. *Georges Demeny | [Fencer].* Available at: https://www.metmuseum.org/art/collection/search/291656 (Accessed: 30 December 2020).

Thompson, K. (1982) *Eisenstein's Ivan the Terrible: A Neoformalist Analysis*. First edition. Princeton, N.J: Princeton University Press.

Tikka, P. (2008) *Enactive Cinema: Simulatorium Eisensteinense*. Aalto University.

Toropova, A. (2017) 'Probing the Heart and Mind of the Viewer: Scientific Studies of Film and Theater Spectators in the Soviet Union, 1917–1936', *Slavic Review*. Cambridge University Press, 76(4), pp. 931–958.

Toulouse, E. and Mourgue, R. (1920) 'Des réactions réspiratoires au cours des projections cinématographiques. 1895 Revue de l'association française de recherché sur l'histoire du cinema' [Respiratory reactions during cinematographic projections. 1895 Review of the French association of research on the history of cinema], 60, pp. 138–144.

Trevarthen, C. (1999) 'Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication', *Musicae scientiae*. SAGE Publications Sage UK: London, England, 3(1_suppl), pp. 155–215.

Trevarthen, C. and Malloch, S. (2017) 'The musical self: Affections for life in a community of sound', in *Handbook of musical identities*. New York, NY, US: Oxford University Press, pp. 155–175. doi: 10.1093/acprof:oso/9780199679485.003.0009.

Trevarthen, C. and Malloch, S. N. (2000) 'The Dance of Wellbeing: Defining the Musical Therapeutic Effect', *Nordisk Tidsskrift for Musikkterapi*. Routledge, 9(2), pp. 3–17. doi: 10.1080/08098130009477996.

Tsivian, Y. (2008) 'Zhest i montazh: eshche raz o russkom stile v rannem kino [Gesture and montage: once again about Russian style in early cinema]', *Kinovedcheskie zapiski*, (88), pp. 65–78.

Tsivian, Y. (2019) *Ivan the Terrible*. Bloomsbury Publishing.

The Metropolitan Museum of Art. *Georges Demeny | [Fencer]*. Available at: https://www.metmuseum.org/art/collection/search/291656 (Accessed: 30 December 2020).

Udine, J. d' (1910) *L'Art et le geste* [Art and Gesture]. Available at: https://gallica.bnf.fr/ark:/12148/bpt6k1265865z (Accessed: 19 December 2020).

Umiltà, M. A., Berchio, C., Sestito, M., Freedberg, D. and Gallese, V. (2012) 'Abstract art and cortical motor activation: an EEG study', *Frontiers in Human Neuroscience*, 6. doi: 10.3389/fnhum.2012.00311.

Umiltà, M. A. and Gallese, V. (2012) 'Abstract art and cortical motor activation: an EEG study', *Frontiers in Human Neuroscience*, p. 9.

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C. and Rizzolatti, G. (2001) 'I Know What You Are Doing: A Neurophysiological Study', *Neuron*, 31(1), pp. 155–165. doi: 10.1016/S0896-6273(01)00337-3.

Vassilieva, J. (2019) 'The Eisenstein-Vygotsky-Luria Collaboration', *Projections*, 13(1), pp. 23–44. doi: 10.3167/proj.2019.130103.

Vaynberg, M. (1957) Film Score for *Letiat Zhuravli*. USSR: Mosfilm.

Vechernaia Moskva – Moi laskovii i nezhnij zver': 5 znamenitih kompozitsii evgeniia doga [Evening Moscow - The Affectionate and Tender Beast [The Hunting Accident]: 5 Famous Compositions of Eugene Doga] (2018). Available at: https://web.archive.org/web/20180201075916/http://vm.ru/news/2014/02/28/moj-laskovij-i-nezhnij-zver-5-znamenitih-kompozitsij-evgeniya-doga-237586.html (Accessed: 3 December 2020).

Vischer, R. (1873) 'On the optical sense of form: A contribution to aesthetics', *Empathy, form, and space: problems in German aesthetics*, 1893, pp. 89–124.

Vogman, E. (2020) 'Le mouvement expressif des plantes. Note sur un projet de film d'Eisenstein ['The expressive movement of plants. Notes on a film project by Eisenstein]', in

Castro, T., Pitrou, P. and M. Rebecchi (eds) *Puissance du végétal et cinéma animiste* [The power of plants and animist cinema]. Les presses du réel.

'Vsevolod Meierhol'd. Ves' etot dzhas' [Vsevolod Meyerhold. All that Jazz] (2016) *Zhurnal Seans* [Journal – Seance]. Available at: https://seance.ru/articles/meyerhold-white-eagle/ (Accessed: 29 December 2020).

Wagner, R. (1913) *Opera and Drama*. Translated by E. Evans. London: WM Reeves.

Wagner, R. (1993) *'The Art-Work of the Future' and Other Works*. Translated by W. A. Ellis. Lincoln: University of Nebraska Press.

Wagner: ring of change (2013) *Socialist Review*. Available at: http://socialistreview.org.uk/380/wagner-ring-change (Accessed: 14 December 2020).

Walker, M. E. (2000) 'Movement and Metaphor: Towards an Embodied Theory of Music Cognition and Hermeneutics', p. 16

Werner, P. D., Swope, A. J. and Heide, F. J. (2006) 'The music experience questionnaire: Development and correlates', *The Journal of Psychology*. Taylor & Francis, 140(4), pp. 329–345.

Wickham, H. (2016) *ggplot2: Elegant graphics for data analysis*. Springer.

Widdis, E. (2017) *Socialist Senses: Film, Feeling, and the Soviet Subject, 1917–1940*. Indiana University Press.

Witmer, B. G. and Singer, M. J. (1998) 'Measuring presence in virtual environments: A presence questionnaire', *Presence*. MIT Press, 7(3), pp. 225–240.

Wolff, P. H. (2002) 'Timing precision and rhythm in developmental dyslexia', *Reading and Writing*. Springer, 15(1–2), pp. 179–206.

Woll, J. (2003) *The Cranes are Flying*. London ; New York : I.B. Tauris.

World Medical Association (2013) 'World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects', *JAMA*, 310(20), pp. 2191–2194. doi: 10.1001/jama.2013.281053.

Worrall, N. (1973) 'Meyerhold's Production of "The Magnificent Cuckold"', *The Drama Review: TDR*. The MIT Press, 17(1), pp. 14–34. doi: 10.2307/1144789.

Wöllner, C., Hammerschmidt, D. and Albrecht, H. (2018) 'Slow motion in films and video clips: Music influences perceived duration and emotion, autonomic physiological activation and pupillary responses', *PLOS ONE*. Public Library of Science, 13(6), p. e0199161. doi: 10.1371/journal.pone.0199161.

Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E. and Evans, A. C. (1996) 'Hearing in the mind's ear: a PET investigation of musical imagery and perception', *Journal of cognitive neuroscience*. MIT Press, 8(1), pp. 29–46.

Zeno. *Volltext, Kulturgeschichte: Bildteil. Lehmann, Lilli: Mein Weg. Leipzig 1913, S. 279-280* [Full text, Cultural History: Pictorial Section. Lehmann, Lilli: My Path]. Available at: http://www.zeno.org/Kulturgeschichte/M/Lehmann,+Lilli/Mein+Weg/II.Teil/Bildteil (Accessed: 14 December 2020).

Zbikowski, L.M. (2002) *Conceptualizing music: Cognitive structure, theory, and analysis*. Oxford University Press on Demand.

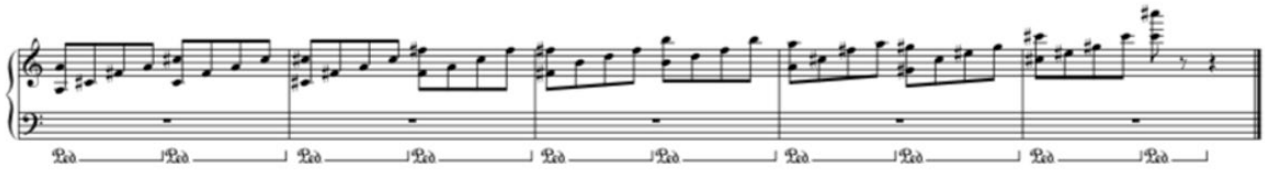
Zuckerkandl, V. (1956) *Sound and Symbol: Music and the External World*. Pantheon.

'80 let so dnia gibelia Vsevoloda Meierhol'da' [80 years since the death of Vsevolod Meyerhold] (2020) *Gazeta.ru*. Available at: https://www.gazeta.ru/culture/photo/80_let_so_smerti_vsevoloda_meierholda.shtml (Accessed: 23 December 2020).

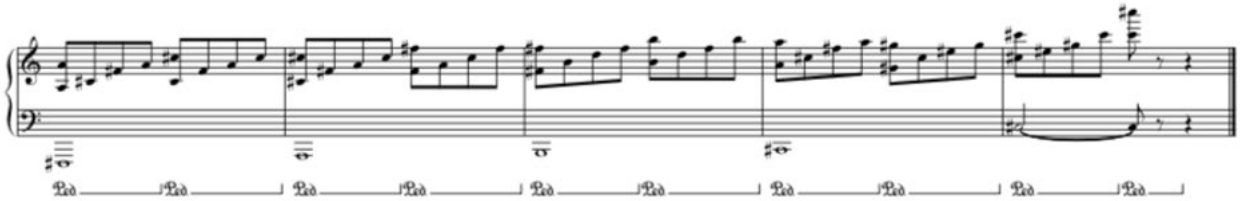
11. Appendix

Partition for Musical Stimuli: Composed by Eduardo Andrade
Intellectual Property of the University of Parma ©
Labels: Contour, Note Pattern, Composition Complexity
♩ = 120, all excerpts and dynamics were played at *mf* as an average.

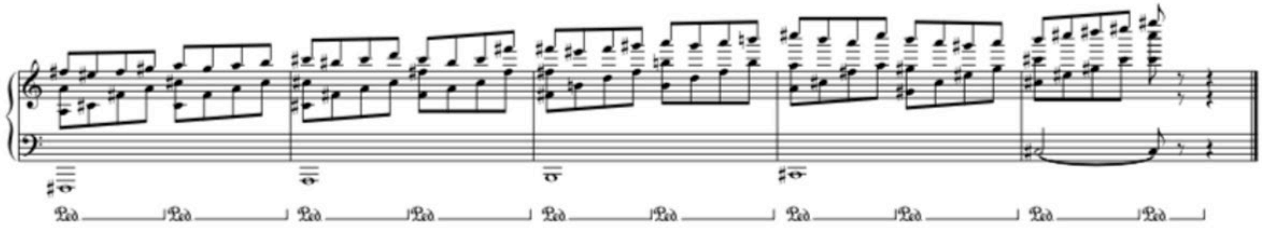
Track 1 | Ascending, Binary, Low



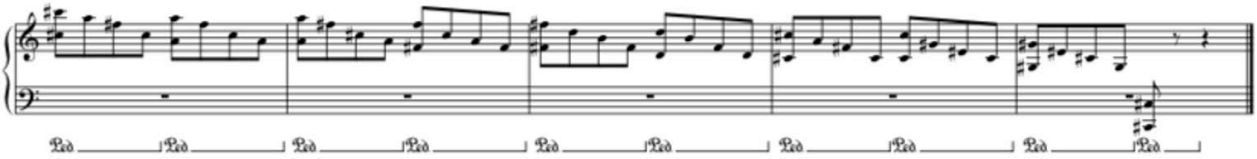
Track 2 | Ascending, Binary, Middle



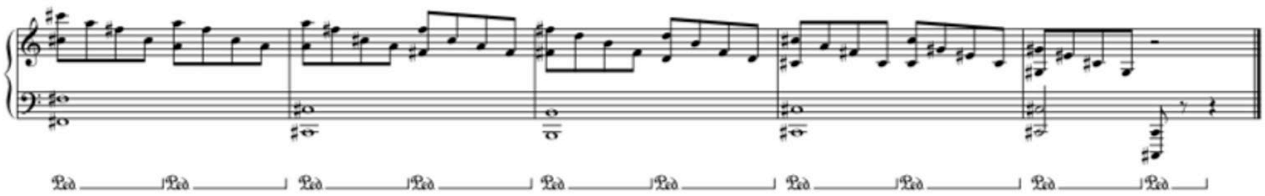
Track 3 | Ascending, Binary, High



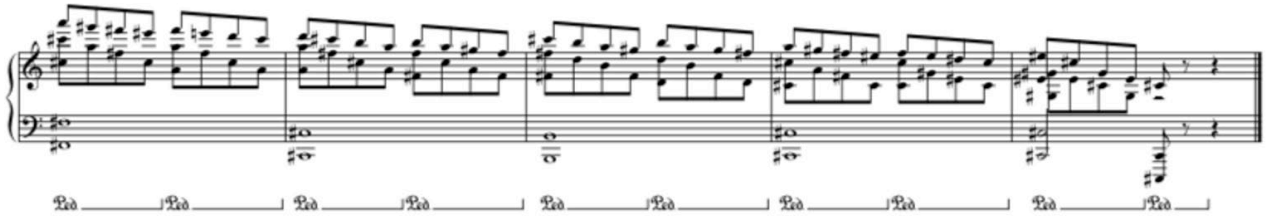
Track 4 | Descending, Binary, Low



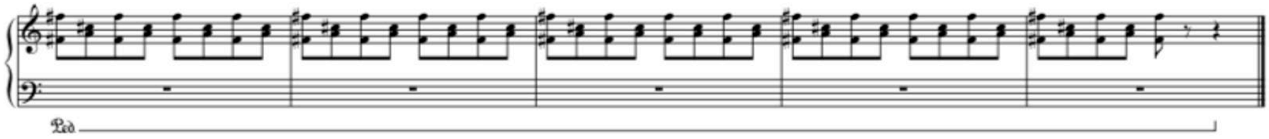
Track 5 | Descending Binary, Middle



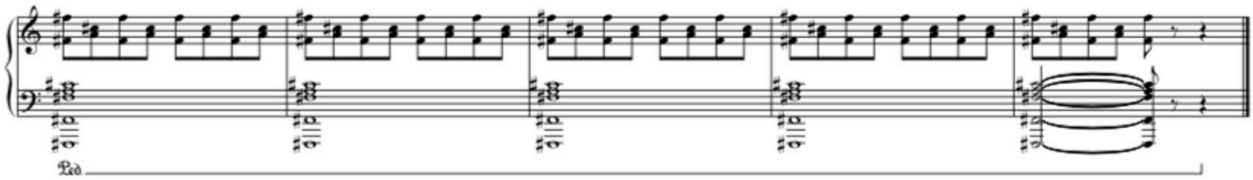
Track 6 | Descending, Binary, High



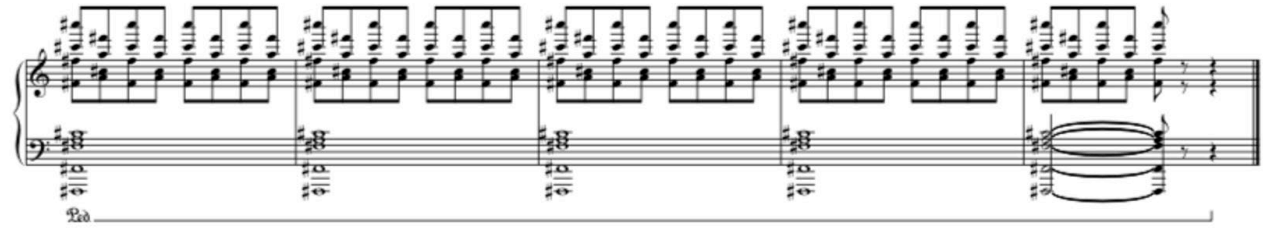
Track 7 | Flat, Binary, Low



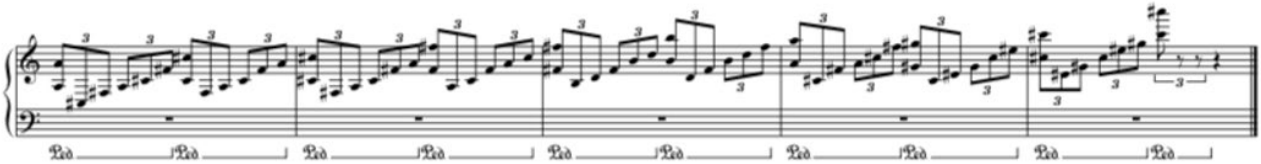
Track 8 | Flat, Binary, Middle



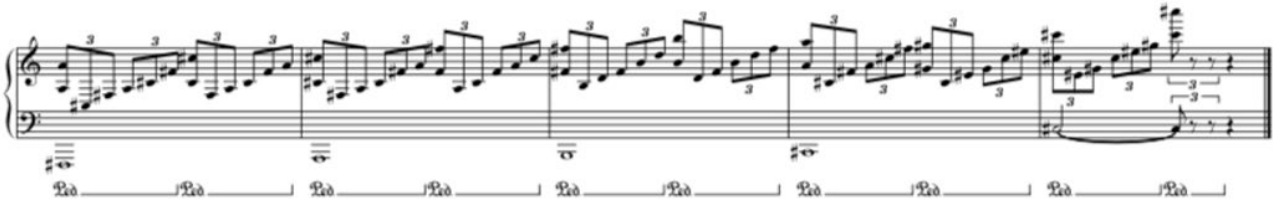
Track 9 | Flat, Binary, High



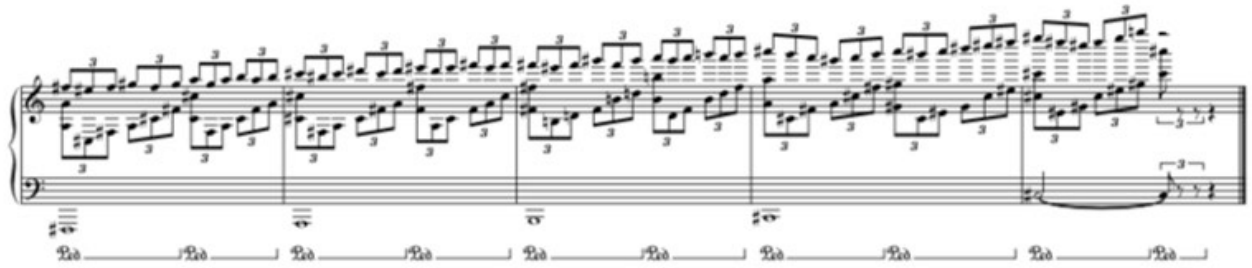
Track 10 | Ascending, Ternary Low,



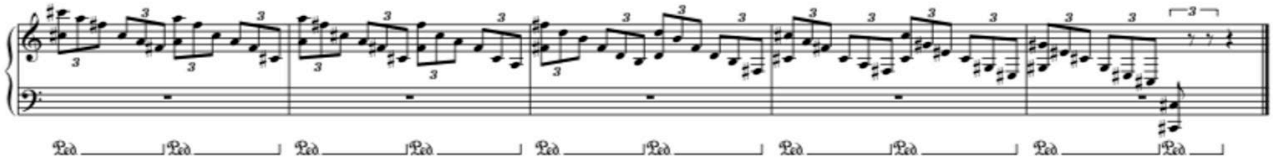
Track 11 | Ascending, Ternary, Middle



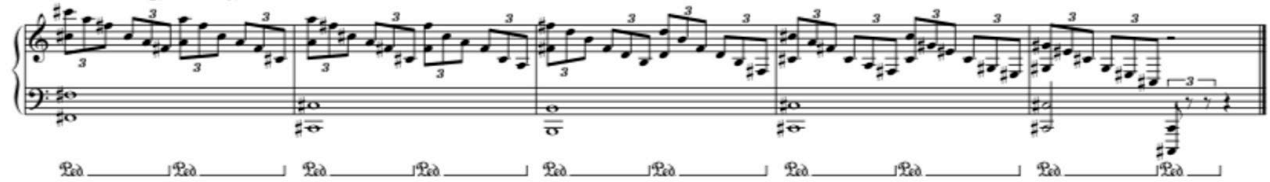
Track 12 | Ascending, Ternary, High



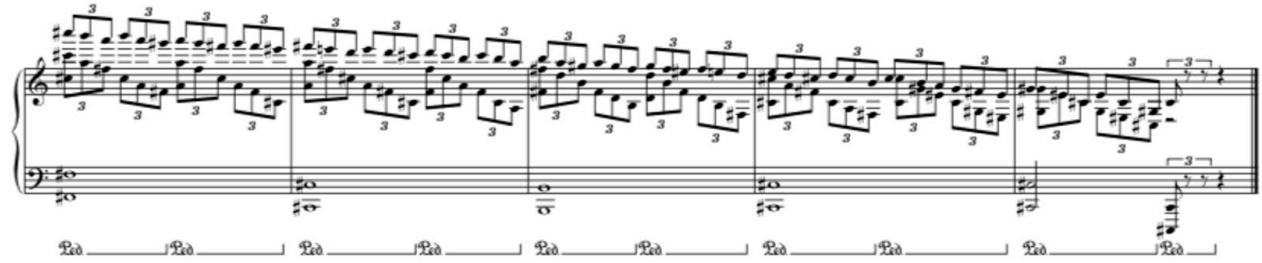
Track 13 | Descending, Ternary, Low



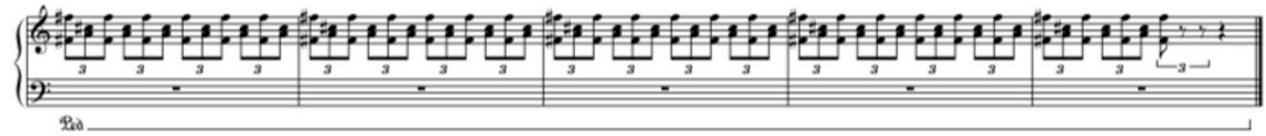
Track 14 | Descending, Ternary, Middle



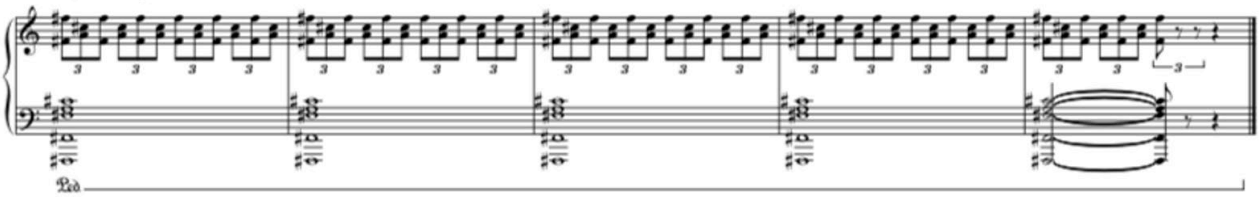
Track 15 | Descending, Ternary, High



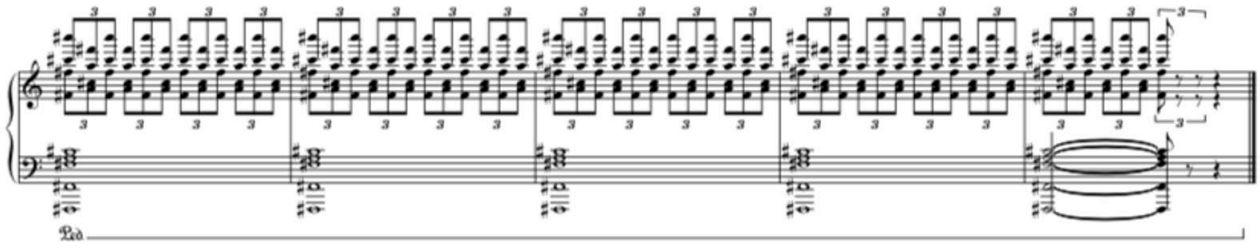
Track 16 | Flat, Ternary, Low



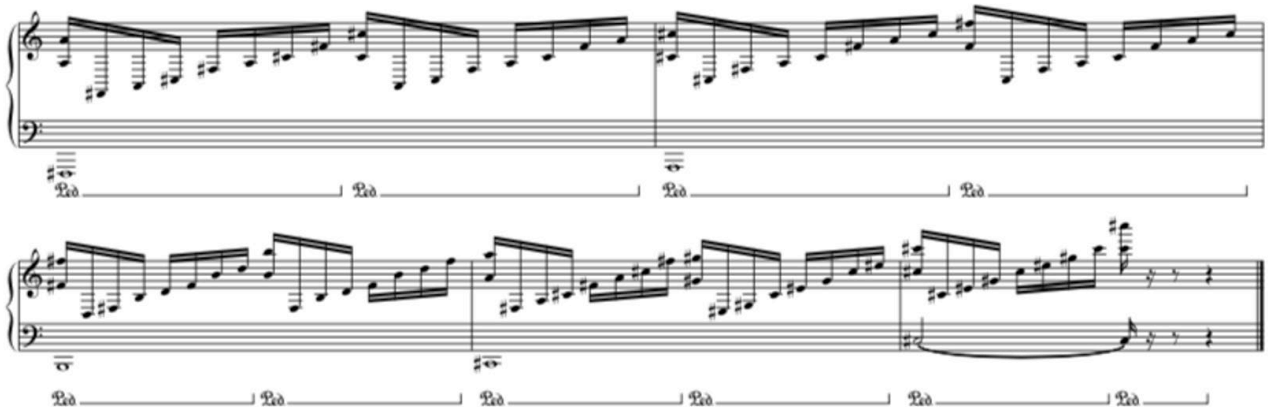
Track 17 | Flat, Ternary, Middle



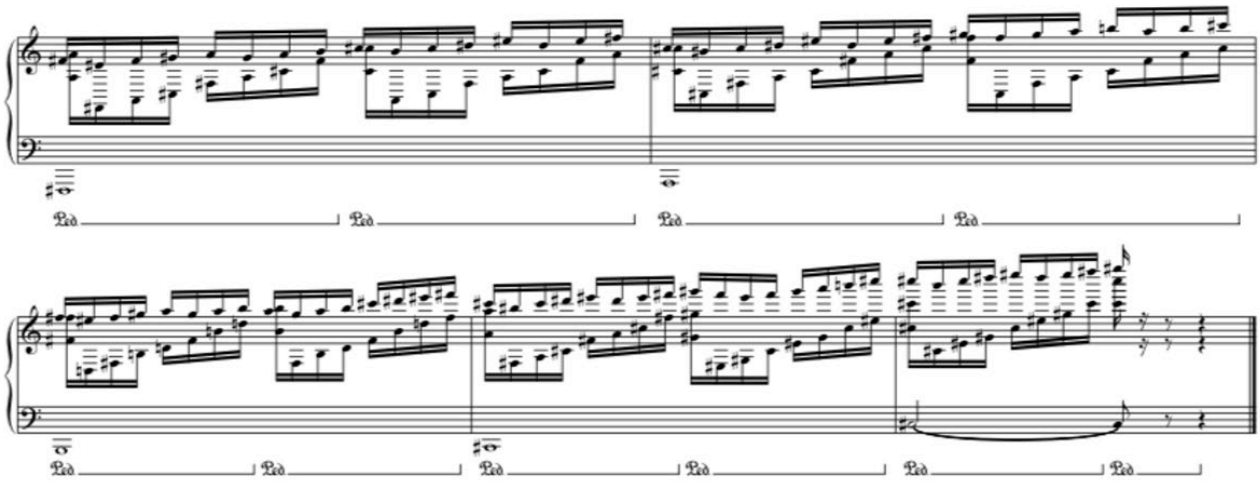
Track 18 | Flat, Ternary, High



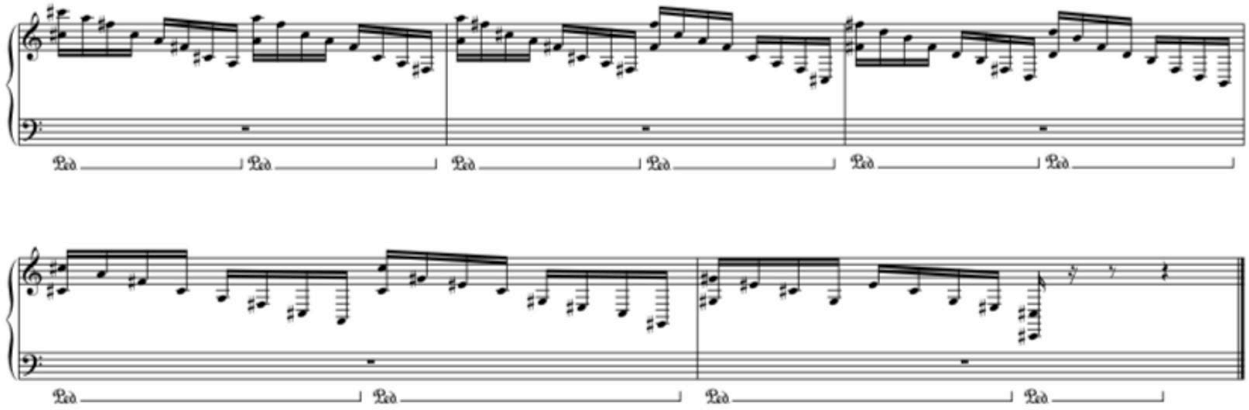
Track 19 | Ascending, Quaternary, Middle



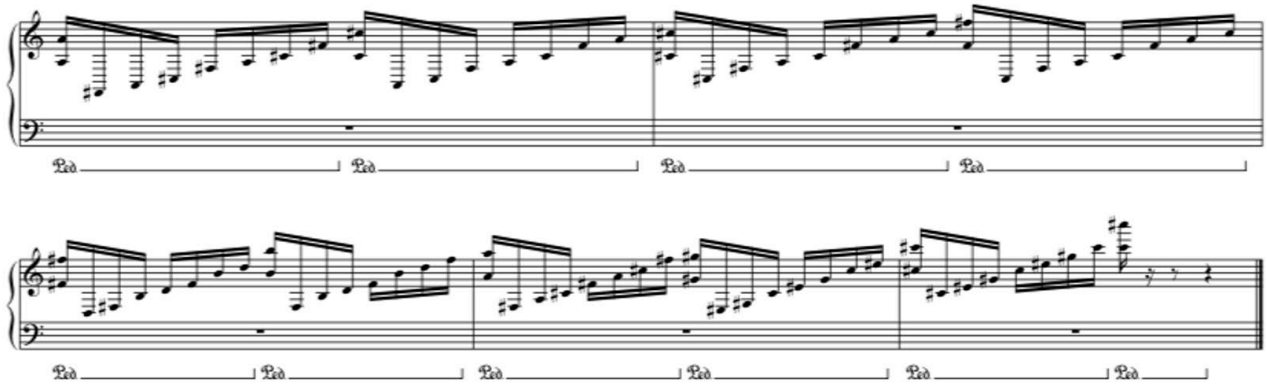
Track 20 | Ascending, Quaternary, High



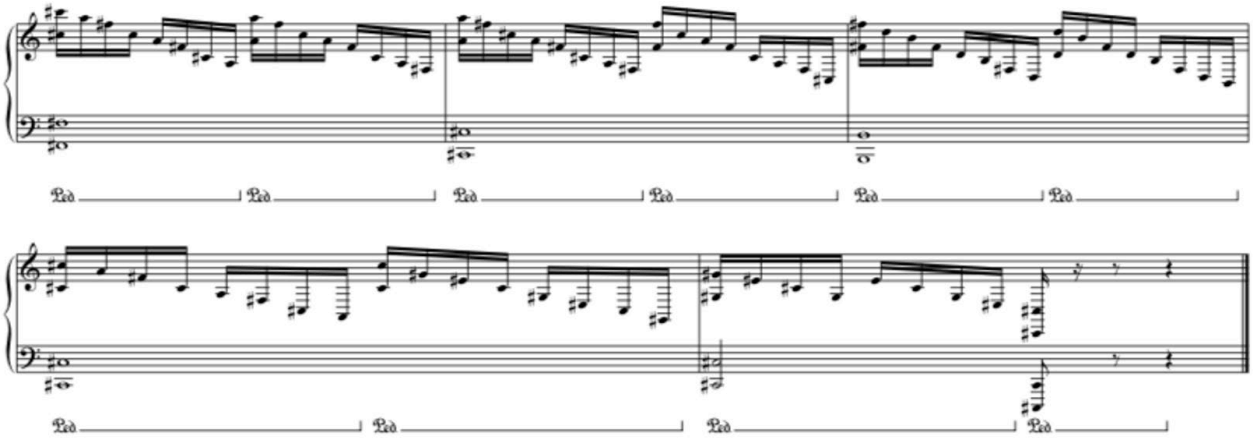
Track 21 | Descending, Quaternary, Low



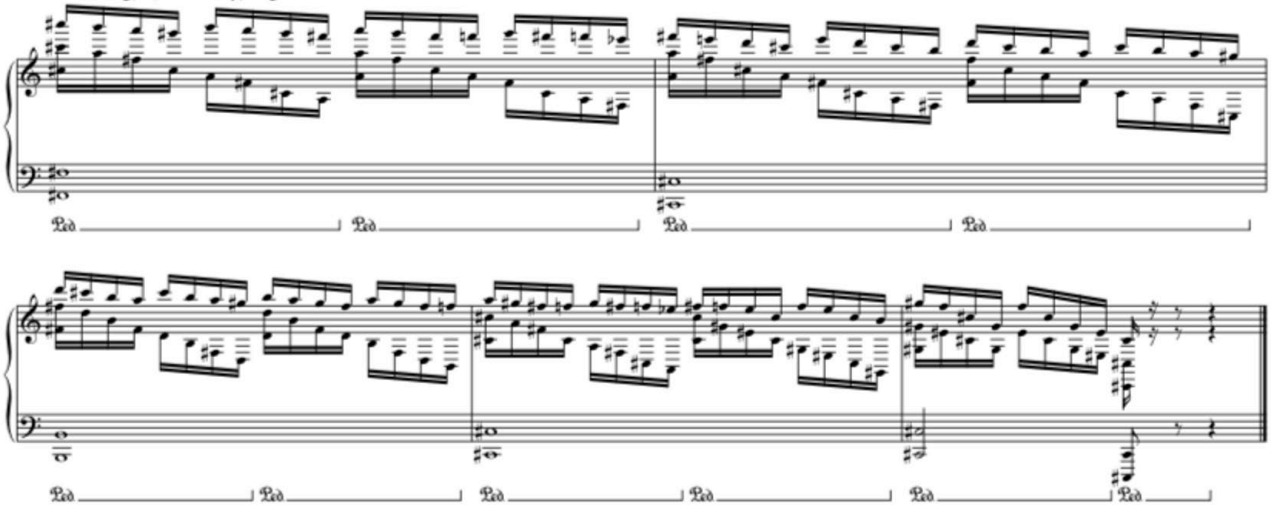
Track 22 | Ascending, Quaternary, Low



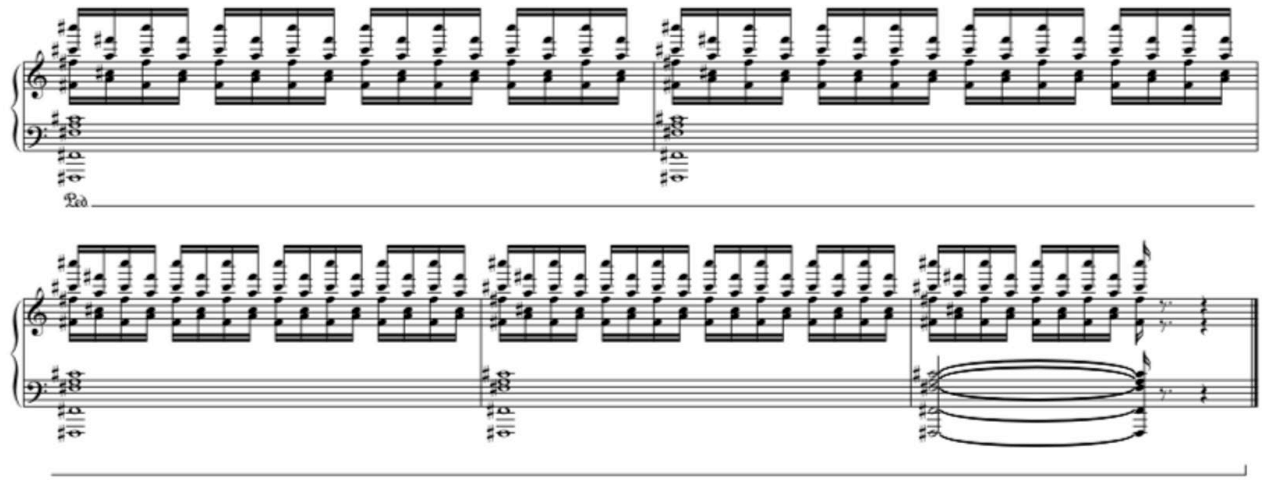
Track 23 | Descending, Quaternary, Middle



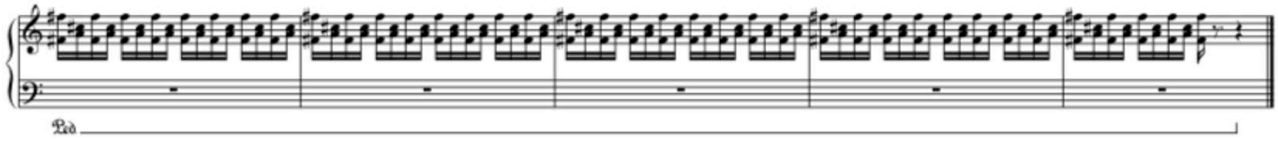
Track 24 | Descending, Quaternary, High



Track 25 | Flat, Quaternary, High



Track 26 | Flat, Quaternary, Low



Track 27 | Flat, Quaternary, Middle