ISSN 3034-3836 (Online)

History and Theory of Music

Original article
UDC 78.01
https://doi.org/10.56620/RM.2026.3.026-037
EDN ZGXXVA

# From "Accompaniment" to "Event": The Evolution of Conceptual Vocabulary in International Research on Audiovisual Synthesis

**Pavel A. Michkov**

*M.I. Glinka Novosibirsk State Conservatory,*
*Novosibirsk, Russian Federation,*
*p.michkov@nsglinka.ru*✉, *https://orcid.org/0000-0001-8227-7190*

***Abstract.*** The article presents a historical and theoretical analysis of the evolution of conceptual vocabulary in international research on audiovisual synthesis and multimedia art. The development of approaches to the methodology of music analysis is traced from utilitarian "accompaniment" in early cinema to the interpretation of the musical component as an autonomous, dramaturgically active "event" in contemporary digital installations. The study focuses on the transformation of methodological approaches that developed in parallel with the increasing complexity of artistic practice. Key events in the chronology of the formation of the interdisciplinary language of science are summarised. The origins of the legitimisation of sonic reality in Francesco Balilla Pratella's Futurist manifesto are examined along with Dziga Vertov's early montage-based and polyphonic concepts of the "radio-ear." Particular attention is drawn to Leonid Sabaneev's first systematic musicological study, which anticipated the foundations of film semiotics through the categories of "aesthetic synchronisation" and the phenomenology of time. Michel Chion's phenomenological approach is examined, in which the category of the "audiovisual contract" and the concept of "added value" provided a means of studying music not as a passive background, but as an active factor in the production of meaning. The trajectory culminates in Nicholas Cook's model, which offers a structured musicological method for investigating multimodal perception. Cook systematically demonstrates that sound cannot be isolated from graphics and context, constructing a semantic typology of musical-visual relationships through the categories of conformance, contest and complementation. Emphasising the cognitive and integrative character of contemporary digital multimedia, which is rooted in Wagner's concept of the *Gesamtkunstwerk*, the article concludes by advocating for a continual renewal of musicology's analytical toolkit.

---

Translated by Thomas Alexander Beavitt, Laboratory for Scientific Translation (https://nauka-perevod.ru/eng).

© Pavel A. Michkov, 2026

***Keywords***: multimedia, conceptual and categorical apparatus, audiovisual synthesis, Leonid Sabaneev, Michel Chion, Nicholas Cook, Louise Harris, polymodal perception, interaction between the arts, dramaturgy of sound event

***For citation:*** Michkov P.A. From "Accompaniment" to "Event": The Evolution of Conceptual Vocabulary in International Research on Audiovisual Synthesis. *Russian Musicology*. 2026, no. 3, pp. 26–37. https://doi.org/10.56620/RM.2026.3.026-037

# От «сопровождения» к «событию»: эволюция понятийного аппарата в зарубежных исследованиях аудиовизуального синтеза

**Павел Александрович Мичков**

*Новосибирская государственная консерватория имени М.И. Глинки,*
*г. Новосибирск, Российская Федерация,*
*p.michkov@nsglinka.ru*✉, *https://orcid.org/0000-0001-8227-7190*

***Аннотация.*** Статья посвящена историко-теоретическому анализу эволюции понятийного аппарата в зарубежных исследованиях аудиовизуального синтеза и мультимедийного искусства. В работе прослеживается путь в осмыслении методологии анализа музыки: от утилитарного «сопровождения» в раннем кинематографе до трактовки музыкального компонента как самостоятельного, драматургически активного «события» в новейших цифровых инсталляциях XXI столетия. В фокусе исследования находится трансформация методологических подходов, развивавшихся параллельно усложнению художественной практики. В статье последовательно раскрываются ключевые события в хронологии формирования междисциплинарного языка науки. Рассматриваются истоки легитимизации звуковой реальности в футуристическом манифесте Франческо Прателлы, а также ранние монтажные и полифонические концепции «радиоуха» Дзиги Вертова. Особое внимание уделено первому системному музыковедческому труду Леонида Сабанеева, предвосхитившему основы киносемиотики через категории «эстетической синхронизации» и феноменологии времени. Рассматривается феноменологический подход Мишеля Шиона, чья категория «аудиовизуального контракта» и концепт «добавленной стоимости» позволили изучать музыку не как пассивный фон, а как активный фактор смыслопроизводства. Итоговой точкой выступает модель Николаса Кука, предложившая структурированный музыковедческий метод исследования полимодального восприятия. Кук последовательно доказывает невозможность изоляции звука от графики и контекста, выстраивая семантическую типологию музыкально-визуальных связей через категории конформности, противоборства и дополнения. Автор подчёркивает когнитивный и интегративный характер современного цифрового мультимедиа, уходящий корнями в вагнеровскую идею *Gesamtkunstwerk*, и доказывает необходимость постоянного обновления аналитического инструментария музыкознания.

***Ключевые слова:*** мультимедиа, понятийно-категориальный аппарат, аудиовизуальный синтез, Леонид Сабанеев, Мишель Шион, Николас Кук, Луиза Харрис, полимодальное восприятие, взаимодействие искусств, драматургия звукового события

## From "Acoustic Glue" to "Chemical Reaction": a Paradigm Shift

Igor Stravinsky's well-known remark, recorded in a 1946 *Musical Digest* interview with the American composer and pianist Ingolf Dahl, and frequently cited in scholarly publications and music criticism, has long been seen as emblematic of the degradation of music in the screen arts. The composer's attitude to this industry is formulated by his famous aphorism: "There is only one real function of film music—namely to feed the composer." [1, p. 4] However, this famous quip is far from Stravinsky's only critical jab at cinema; in addition, he laments that, on the screen, music "…has got to bridge holes; it has got to fill the emptiness of the screen and supply the loudspeakers with more or less pleasant sounds." [Ibid., p. 35] In Russian translation, Stravinsky's references to "bridging holes" and "filling the emptiness of the screen" became established in music criticism and Stravinsky scholarship in the condensed expression *zapolnitel' pustot* ("gap filler"), subsequently encapsulated in the phrase "acoustic glue."

These outrageous lines instantly became cult classics and are still widely circulated in numerous critical sources. Less often noted, however, is another passage from the same interview, in which the composer formulates a demanding aesthetic ideal for works of the screen arts through the following conceptually productive metaphor: "…my ideal is the chemical reaction where a new entity, a third body, results from uniting two different but equally important elements: music and drama. This is not the chemical mixture where, as in the films, to the preordained whole just the ingredient of music is added, resulting in nothing either new or creative." [Ibid.]

In this radical manifesto directed against Hollywood's commercial film industry, the composer envisages the future in terms of an autonomous audiovisual architecture in which music is created according to different principles, rather than being subordinated to the elementary logic of sonically duplicating the plot. This historical trajectory—from early cinema to contemporary multimedia installations—encompasses not merely a century of development, but a fundamental shift in the understanding of audiovisual synthesis. Having moved beyond the illustrative stage, contemporary audiovisual art has arrived at an almost complete autonomy of its visual and sonic media layers. Stravinsky's prophetic vision, which captured the profound ontological crisis of the future of multimedia[1] in the middle of the last century, crystallised over the years. In parallel, musicology developed a distinct area of research devoted first to music in the screen arts and later to music in multimedia art. This field developed in parallel with the evolution of its conceptual apparatus and scholarly terminology. The musical component of multimedia experiments is difficult to study using established methods, since such works require a different analytical toolkit capable of deciphering the complex relationships between sound, space and the digital image.

[1] The term "multimedia," which had previously been applied neither to experiments combining light and music nor to works of cinema, emerged in Western culture only in August 1966. It was coined by musician Bobb Goldstein, who used the neologism to advertise his avant-garde light and music show in New York. The first printed mention was recorded on 10th August 1966 in a review printed in the American magazine *Variety*. The term then entered academic musicological literature at the turn of the 1960s and 1970s mainly to describe rock concerts and instrumental theatre works.

## Methodological Contours of Audiovisual Synthesis: in Search of a New Language

Comparing the focus of art history research in the last century and the present, one cannot help but notice a number of facts accompanying the study of music in the field of audiovisual and multimedia compositions. Thus, despite the abundance of contemporary works of a highly technical nature, today another aspiration is clearly visible—both on the part of researchers and on the part of the creators of "multimedia" [2] compositions themselves. Their desire to move beyond the purely craft-based and technical dimensions of multimedia largely determines the new character of the theory of multimedia art. Such changes in the search for new compositional techniques and alternative editing methods, together with new forms of collaboration with directors on visual effects and new approaches to analysing the perception of a composition, are likely to be closely connected with the broader state of audiovisual practice and digital culture as a whole.

A substantial body of scholarship, both Russian and international, has addressed the complexity and diversity of the individual branches that make up the tree of contemporary media art in the 21st century. There is probably no need here to enumerate the questions and lines of enquiry currently occupying researchers, since this field is too diverse to summarise adequately and would merit a separate publication. Nevertheless, it becomes necessary to recall this phenomenon, because it is precisely this that gives rise to the need for a methodological explanation, which, in turn, is only possible on a sufficiently broad basis. Hence the search for a theoretical foundation, prompting different authors to draw on interdisciplinary approaches and semiotic systems regarded as most productive at the time of a given digital work's creation. At the same time, a full disclosure of the methodological foundations of multimedia concepts—including, in particular, the process of forming one's own theoretical base and conceptual apparatus—presents considerable difficulties.

It is worth acknowledging that contemporary new-media theory today contains virtually no research conducted within the closed disciplinary boundaries of art studies alone. Most often, the platform for such work is a ramified integration of approaches drawn from related humanities and technical disciplines. Contemporary theoretical works created in recent decades that pursue general methodological objectives are characterised by a natural intersection of various research vectors, which is expressed in the inherently interdisciplinary nature of the phenomenon under consideration. This convergence arises from what is taken to be the principal task of this type of theory: providing a systematic rationale for new and unconventional forms of interactive art. The need to account for the unexplored therefore calls for an interdisciplinary approach, in the hope that it can demonstrate the internal logic of novelty, however unfamiliar that novelty may appear to the reader.

This seems to explain the frequent attempts to theorise the latest multimedia compositions, whose technological diversity in the digital age calls for the construction of clear conceptual models. In this regard, the basic concepts are formed at the intersection of art history, semiotics, film theory and computer science.

[2] As most commonly used in Russian and international scholarship, the meaning of the term "multimedia" appears as informed by the adduced discussion. In particular, Valentina Kholopova in her article about the work of Igor Kefalidi gives a description of "multimedia": "…among these 'media' there may be graphics, verbal text, sound, video images, animation and others, with which a person works with the help of a complex of hardware and software, synthesising them into a single information field." [2, p. 11]

The task of a deep theoretical justification in this case consists in revealing the internal artistic persuasiveness and logic of the proposed interactive method. Here it seems indicative to turn to the structural analysis of the screen image, where the key distinction becomes the division between the physical nature of the digital signal and the "audiovisual object" as a semiotic unit. Thus, the principal problem in developing an appropriate methodology consists in describing such an object, which functions at the interface between the visual continuum and the sonic mass.

In the present work, it becomes necessary to outline the chronological boundaries of the phenomenon under consideration. In this way, the object of the study becomes the conceptual apparatus of music analysis in audiovisual arts, the visual series of which is created on the basis of modern technologies. The starting point of the research is the evolution of screen sound—from the practice of pianist and the creation of original film scores to the interactive multimedia forms of the 21st century. Musical theoretical analysis in the West developed in several waves following this process, which consistently move away from an understanding of classical cinema to the theory of digital multimedia.

## Avant-garde Origins: from Pratella to Vertov

The historical origins of this conceptual evolution go back to the beginning of the 20th century, representing an era of a radical break with the classical academic tradition. The first theoretical outlines of the future sound-visual synthesis were born at the intersection of futurism and technological progress. Thus, already in 1910, the Italian composer and theorist Francesco Balilla Pratella in his *Manifesto of the Futurist Musicians*[3] declared the need to expand the sound boundaries of art. Comparing different national schools of composition, Pratella notes that while "in France Claude Debussy…floats in…a tranquil lake of…transparent harmonies," and "in Russia Modest Mussorgsky, renewed through the soul of Nikolai Rimsky-Korsakov, through the grafting of a primitive national element… causes tradition to be forgotten,"[4] Italian Futurism proposes a fundamentally different direction for the development of art.

In the same text, Pratella criticises academic institutions, calling instead for "drawing inspiration and aesthetics from nature, through all its phenomena—real, human and superhuman."[5] These phenomena implied a new urban and industrial environment. Pratella opens the door to non-harmonic sounds (noises and ambient sounds), which were previously excluded from the space of art and considered "anti-musical." The manifesto thus legitimised the sound reality of the new world as a fully-fledged aesthetic object.

These ideas would later prompt the Soviet film director Dziga Vertov to develop a systematic methodology for working with musical material in the screen arts. Even in the pre-sound period, in his early manifestos and theoretical notes (beginning with the article *WE: Variant of a Manifesto*, 1922), Vertov effectively transferred the laws of musical literacy to the plane of film editing. According to his theoretical concept, recorded sounds and visual segments were likened to notes, while the structure of the work was built on "intervals"—rhythmic, dynamic and semantic transitions between them.

[3] Pratella F. B. *Manifesto dei musicisti futuristi.* Milan, 1910. 4 p. (Manifesti futuristi / a cura di L. Caruso; 14). URL: https://www.memofonte.it/files/Progetti/Futurismo/Manifesti/I/14.pdf (accessed: 10.08.2026).

[4] Ibid., pp. 1–2.

[5] Ibid., p. 3.

In his later writings[6] on the concepts of the "radio-eye" and "radio-ear," Vertov would develop a methodology in which temporal sound elements cease to function as an illustrative background. These are organised into an autonomous, polyphonic audiovisual fabric. Thus, Vertov's theoretical legacy laid the foundation for the structuring of sound in film studies, transforming the intuitive concepts of the Futurists into a rigorous editing system.

## Leonid Sabaneev's Theory of Sound-Visual Synthesis

The further evolution of theoretical thought in the 1930s was marked by the publication of fundamental works that summarised the international experience of sound cinema. The key figure of this stage was the musicologist, music critic, and composer Leonid Leonidovich Sabaneev, who published the monograph *Music for the Films* in 1935. [4] Although written in Russian, the book was first published in English translation by Samuel William Pring and published in London by *Sir Isaac Pitman & Sons*. Representing a comprehensive methodological study of sound-visual synthesis, the monograph was written at a time when sound cinema had existed for less than ten years. Sabaneev was perhaps the first to transfer the strict categories of academic musicology into the space of screen arts, thus laying the foundations for future theories of multimedia. The author's principal concept is built around the phenomenology of screen time and musical meter. He notes that visual editing (changing frames) has a discrete nature, while music is characterised by continuous duration. When superimposed on the visual sequence, musical tempo and rhythm can slow or accelerate the viewer's subjective perception of the action. Sabaneev introduces the concept of "aesthetic synchronisation," arguing that the mechanical coincidence of a strong beat with a visual accent is detrimental to art because it reduces music to a kind of poster: "Aesthetic synchronization has nothing in common with the crude mechanical coincidence of the strong beat and the visual accent; a genuine synthesis arises only when the music preserves its own internal logic and structural integrity, developing parallel to the frame rather than servilely imitating it." [Ibid., p. 22]

In laying the foundations of the semiotics of audiovisual works, Sabaneev anticipates Michel Chion's famous differentiation into "diegetic" and "non-diegetic" sound. Sabaneev divides film music into two methodological categories: background or illustrative-decorative music—a soundtrack used to create atmosphere and to "glue together" the joins between shots—and dramaturgically active music, which becomes an independent agent within the action. The latter may contradict what is shown on screen, reveal a character's hidden thoughts or foreshadow events that have yet to appear. This closely anticipates what Nicholas Cook would, some sixty years later, describe as the models of "complementation" and "contest."

The principal methodological thrust of Sabaneev's work partly anticipates Stravinsky's later manifesto. Sabaneev rejects the utilitarian approach in which music becomes the "handmaiden" of the image. He insists that even within a three-minute film episode, the music must maintain the internal logic of the musical form. At the same time, he does not deny that "...in general, music needs to understand that in cinema it should almost always remain in the background: it acts, so to speak, as a tonal figuration—the 'left hand' to the melody on

[6] The speech is, in particular, about the report *Kinoglaz, radioglaz and the so-called documentary*, [3] presented by Vertov to the Main Political Education Directorate of the RSFSR in 1929.

screen—and things go wrong when that left hand starts pushing into the foreground and obscuring the melody." [Ibid., p. 17]

## "Audiovisual Contract" by Michel Chion: Sound as a Generator of Meanings

More than half a century later, the methodological line outlined in Sabaneev's work had evolved into Chion's large-scale phenomenological concept,[7] which was set out in his work *Audio-Vision: Sound on Screen.* [5] Chion decisively moved the analysis of cinematic sound beyond the limits of the established framework by introducing the "audiovisual contract" into academic discourse. His book was first published in French by the renowned Parisian publishing house *Éditions Nathan* in 1990, translated into English in 1994 by Claudia Gorbman, an American film scholar, Ph.D., professor emeritus at the University of Washington in Tacoma, and published in New York by the university press *Columbia University Press* in the same year.

A notable feature of this study is the author's focus on the sound component of cinema through the introduction of the concept of "audiovisual contract." [Ibid., p. 3] This concept occupies a central place in the architectonics of the work: the researcher not only opens his or her work with it, placing it in the title of the first part, but also develops a deep theoretical analysis on its basis. The author advances the idea that the perception of multimedia art is based on a kind of agreement, or "symbolic pact," [Ibid., p. 222] between the creative team of director and composer and the viewer. Within the framework of this pact, auditory and visual stimuli do not duplicate, but mutually determine and qualitatively transform each other, giving birth to fundamentally new meanings. A kind of synergetic fusion of acoustic and visual series arises. This kind of integration forms a single "transsensory" whole, [Ibid., p. 136] within which sound qualitatively transforms the semantic content of the frame, while the image, in turn, determines the perception of auditory material.

Chion's earlier study of the phenomenology of sound perception in the screen arts has been extensively explored within Russian musicological scholarship. The attention of contemporary Western musicology is increasingly focused on those latent aspects of Chion's concept that for a long time remained on the periphery of research and constituted a kind of "behind-the-scenes" space of his theory. Although primarily received as a scholar of the phenomenology of sound perception, Chion implicitly embedded in *Audio-Vision* a new analytical toolkit for musicology, whose potential is only now being realised in the latest research on media art. Authors note that Chion took the analysis of film music far beyond a simple analysis of scores or illustrative functions. This has prompted a reassessment of the musicological propositions that underpin the French composer's conception. Thus, Sara Pinheiro[8] notes that Chion "does not theorize about sound per se but about elements that happen to be audible: voice, speech, text, music."

[7] Michel Chion (b. 1947) is a French film critic and composer, a representative of musique concrète. He began his career in 1970 as assistant to Pierre Schaeffer at the Groupe des Recherches Musicaux (GRM).

[8] Sara Pinheiro (b. 1985) is a sound engineer, teacher and sound designer. Pinheiro's research interests include sound theory in film production. She received her degree in cinematography in Lisbon in 2008, after which she continued her studies in the Netherlands. She was awarded Doctor of Philosophy (Ph.D.) by the School of Music and Media, Bangor University, Wales. Since 2013, she has been actively teaching in the Czech Republic, serving as a permanent lecturer at the Centre for Audiovisual Research at the renowned Prague film school FAMU and as a guest lecturer at the Royal Conservatory of Music in The Hague.

[6, p. 142] His focus is on the impact of sound on the overall poetics of cinema, rather than "sound as a language in itself. His attention belongs to the filmmakers rather than the sound-makers." [Ibid., p. 142] Despite this, Chion's book remains the most frequently cited source for the fundamental conceptual framework used to analyse the interaction between sound and image in the screen arts and multimedia. Nick Poulakis draws attention to this point,[9] emphasising that Chion's work represented a major milestone in the field because it "has been the first book that systematically theorized the field of 'film sound/film music studies'." [7, p. 458] Poulakis states that Shion's text "moves the discussion away from 'synaesthesia'." [Ibid., p. 459] Instead, the category of "audiovisual contract" is introduced, which is understood as a kind of "symbolic contract that the audio-viewer enters into, consenting to think of sound and image as forming a single entity." [Ibid.] Within the framework of this model, music is explored not as a passive accompaniment to the image, but as a dramatically active element, an event that radically recodes the visual sequence. The key analytical tool here is the concept of "added value," which reveals "the expressive and informative value with which a sound enriches a given image." [Ibid.] Chion's concept radically restructured the conceptual apparatus of multimedia researchers. It can be said to have transformed researchers' intuitive explorations into a rigorous, systematic methodology for analysing the audiovisual arts.

Pierre Schaeffer's theory of the "sound object" became an important instrument of such systematisation, proposing that sound be studied in isolation from its source and cultural context and thereby laying the theoretical foundations for analysing contemporary digital, collage-based and multi-medium compositions. However, this method was not flawless. Schaeffer's peculiar methodological eclecticism is remarked on by Józef Kon in his 1981 work. Analysing the concept of concrete music, the researcher explains this eclecticism as a forced necessity—the search for a theory "aimed at substantiating new, non-traditional forms of musical art." [8, p. 5] Moreover, Kon advances the critical argument that, in musique concrète, "elaborate theoretical argumentation is intended <…> to compensate for the insufficient artistic persuasiveness of the actual result produced by the proposed <…> compositional method." [Ibid., p. 7] This thesis could well apply today to the contemporary terminological system of multimedia, where complex digital theory also often outstrips or justifies the imperfections of artistic practice itself.

### Multimedia as an Interactive "Event"

The period of the late 20th century was marked by the mature formation of a theory of audiovisual synthesis into an independent scientific discipline. Whereas Michel Chion's work during this period was concerned primarily with systematising the classical audiovisuality of cinema, by the end of the decade it had become necessary to extend this approach to new digital formats. A conceptual answer appeared in the form of a book by the famous British musicologist, professor at Oxford and Cambridge Universities, member of the British Academy (FBA) Nicholas Cook *Analysing*

[9] Nick Poulakis is a Greek researcher, ethnomusicologist, cultural scientist, Ph.D., and a staff member of the Laboratory of Ethnomusicology and Anthropology at the Kapodistrian University of Athens. Author of a dissertation on music in Greek cinema of the 1960s. He taught anthropology of music, documentary filmmaking, ethnographic film and music informatics at universities in Greece. He studies film music as a cultural artefact, explores Balkan folklore and creates ethnographic films. Member of the International Music and Media Research Group and the editorial board of the journal *Polyphonia*.

*Musical Multimedia*, [9] published in 1998 by *Clarendon Press*. In it, the author, drawing on Chion's findings, proposes a structured approach to the analysis of the interaction of sound and image, henceforth considering them as equal components of meaning production. Cook introduces a theoretical model of semantic interaction (conformity, opposition, complementation) that describes in detail the types of connections between auditory and visual texts.

Notably, the conceptual framework employed by the author takes account of the distinctive features of multimedia compositions, which are inseparably linked to multimodal perception, and can provide a foundation for a methodology of music analysis applicable to works of this kind. Cook consistently argues that, in the contemporary context, sound cannot be analysed in isolation from graphics, text and context, since these elements function together as a unified semantic construct.

By examining the applicability of synaesthesia to multimedia analysis and the significance of the relevant theoretical framework, the author defines the central concern of the study as the methodological basis for analysing music in multimedia compositions. Such questions have generally remained at the margins of academic enquiry and have appeared only rarely—if at all—in studies devoted to multimedia art.

The study maintains a balanced division of the work into two sections, the first of which is devoted to the theoretical basis used by the author as an analytical tool, while the second comprises a series of analytical studies in which this particular concept is applied. As a general outline, the extremely clear organisational structure presents the reader with a functional and concise method of analysis with examples of its practical application.

The idea formulated in the text of the work that "the identification of word with denotation and music with connotation suggests the kind of layered, non-competitive relationship" [Ibid., p. 119] gradually becomes the catalyst for the main idea that leads the author to his own analytical model. In general, its main characteristics are outlined in the paradigm of connections: identifying similarities and analysing differences. The essence of the study of connections consists in checking whether media components duplicate the same meaning across different modalities of perception, or whether these elements convey inconsistent messages. In the latter case, the perception of the listener-viewer is transformed into a more complex interpretative process. According to Cook's typology, the interaction of arts in multimedia is structured within a single space and can be reduced to three key models. The first model operates with the concept of *conformity*, in which the auditory and visual components duplicate the same message. At the opposite pole is the situation of *confrontation*, where the meanings of different media come into direct conflict with each other. An intermediate position is occupied by *complementary* relationships, which are characterised by a dynamic balance; here, the connection between the components is neither absolutely identical nor clearly conflicting.

In this regard, the anthology edited by Randall Packer and Ken Jordan [10] is noteworthy. In it, the authors propose relying on a strict differentiation that depends on what exactly is understood by multimedia composition. The researchers challenge the myth that multimedia compositions became possible only with the invention of the personal computer. The approach to multimedia in this publication is not technological, but rather conceptual and cognitive: the computer is seen as merely the best means, offering a platform for the implementation of long-existing ideas and concepts. It is a kind of "meta-medium," [11, p. 126] capable of physically embodying the long-standing aesthetic quests of humanity. British artist Roy Ascott puts it even more clearly: "Multimedia, in my view, is not an invention but an ongoing discovery of how the mind and the

universes it imagines (or vice-versa, depending) fit together and interact." [12, p. 311]

In analysing the anthology's content, Packer and Jordan identify five universal pillars of multimedia: integration, interactivity, hypermedia, immersion, and narrativity. Tracing a direct connection between Richard Wagner's operatic concept of the *Gesamtkunstwerk* and contemporary virtual environments, they show that the ideal of total viewer immersion long predates the advent of microchips. Citing the happenings of John Cage and Allan Kaprow as examples, the researchers point to early models of user interactivity that, even at that stage, already called for an updated methodological framework.

In recent research in the field of multimedia arts, particularly in the methodology of audio analysis, there has been a gradual shift away from classical musicological analysis towards interdisciplinary approaches involving the psychology of perception, information technology, and cognitive science. Music analysis based on the models of Nicholas Cook or Michel Chion has tended to focus primarily on semantic correspondences—how a pre-existing sound comments on or complements a pre-existing image. In her monograph, Louise Harris[10] proposes shifting the focus to the level of microstructures—both in the field of musical expressive means and in the sphere of generating a visual series. Harris identifies a distinctive property of multimedia compositions, which she terms cross-modal plasticity, referring to the capacity of auditory and visual parameters to flow into one another, be mutually converted and undergo algorithmic exchange in real time. In contemporary multimedia compositions such as *Syn* by Ryoichi Kurokawa, *Fundaments and variants* by Yannis Kyriakides, as well as *Auroculis* by Harris herself, music becomes a generative space for the formation of a visual and eventful series. The pitch, texture density, timbral richness, or rhythmic pattern of an audio signal can be used to directly control the behaviour of digital pixels. According to Harris's concept, the boundaries between the auditory and visual layers in the digital environment are completely blurred, turning into mobile spaces of mutual translation and transmutation. [13]

The synthesis of cross-modal plasticity and the subjective factor serves as a basis for Harris' formulation of a key methodological conclusion: an interactive audiovisual work is not a static text, but a perceptual process. Music in such systems encodes space, while the subject decodes it through his or her involvement.

## Conclusion

The development of a methodology for studying multimedia compositions is structured by an interdisciplinary approach. The evolution of theoretical models and artistic practices in multimedia art, as traced in international scholarship, reveals a transformation that has come to define the contemporary branch of musicology devoted to multimedia compositions. Classical concepts of audiovisual counterpoint and multisensory perception, which laid the foundation for the study of the interaction of sound and image in the 20th century, are now faced by the need for a gradual rethinking of the categorical apparatus. Traditional analytical models, most clearly represented in the works of

[10] Louise Harris (b. 1983)—professor at the University of Glasgow, media composer, sound artist. She received her higher education at Oxford University and defended her doctoral dissertation at the University of Sheffield (2010). Her research and creative interests include electronic music, computer graphics, fixed media, live performances and large-scale spatial installations. Harris has gained international recognition as a researcher of the phenomenon of expanded audiovisual formats and cross-modal interaction. Her musical and multimedia works are regularly exhibited at major international festivals in the UK, Europe, the USA and Canada.

Nicholas Cook and Michel Chion, were initially formed in the space of cinema and early screen arts, where musical and visual components existed as autonomous, complete and static media realities. Within this approach, the research focus inevitably shifted towards semantic, illustrative or contrapuntal correspondences.

It is noteworthy that the presented evolution of the conceptual apparatus in foreign studies of audiovisual synthesis demonstrates a conceptual correlation with similar theoretical processes in Russian musicology. In the Russian scholarly tradition, the foundation of the study of screen arts has historically relied on the classical theories of montage by Sergei Eisenstein, the concept of "visible music" and the phenomenon of synaesthesia (polysensory perception), developed in the works of Bulat Galeev and representatives of the Prometheus research centre. In these conceptions, sound and image were treated as autonomous components, while their synthesis was analysed within the confines of the frame.

The current state of Russian musicology records the transformation of methodological orientations. In research on new media—ranging from detailed analyses of music-computing technologies to the semiotics of virtual environments—an intermedial approach is becoming increasingly prominent. Traditional musicological categories (such as "composition," "text," and "form") in Russia, as well as abroad, are showing potential for substantive expansion. The study of stable audiovisual counterpoint is being replaced by the analysis of open, processual and interactive systems (for example, in generative art or multimedia performances). Thus, both Russian and foreign research paradigms are moving synchronously from a text-centric analysis of isolated media to a dynamic, interdisciplinary understanding of the digital environment.

## References


1. Stravinsky I. Igor Stravinsky on Film Music. *Musical Digest*. 1946, vol. 28, no. 1, pp. 4–5, 35–36.
2. Kholopova V.N. The concept of "multimedia" in the music of Igor Kefalidi. *Zhurnal Obshchestva teorii muzyki* (Journal of the Music Theory Society). 2019, no. 4, pp. 10–17. (In Russ.)
3. Vertov D. Kinoglaz, radioglaz and the so-called "documentalism." *Proletarskoe kino* (Proletarian cinema). 1931, no. 4, pp. 12–15. (In Russ.)
4. Sabaneev L. *Music for the Films: A Handbook for Composers and Conductors*. Transl. by S.W. Pring. London: Sir I. Pitman & Sons, 1935. 128 p.
5. Chion M. *Audio-Vision: Sound on Screen*. New York: Columbia University Press, 2019. 270 p.
6. Pinheiro S. Revisiting the Canon of Sound Theory. *Iluminace*. 2021, vol. 33, no. 1, pp. 141–147. https://doi.org/10.58193/ilu.1699
7. Poulakis N. Review of Audio-Vision: Sound on Screen (Second Edition). *CINEJ Cinema Journal*. 2020, vol. 8, no. 2, pp. 488–492. https://doi.org/10.5195/cinej.2020.289
8. Kon Yu.G. On the methodological foundations of Pierre Schaeffer's theoretical views. *Music Journal of Northern Europe*. 2026, vol. 45, no. 1, pp. 1–28. (In Russ.) https://doi.org/10.61908/2413-0486.2026.45.1.1-28
9. Cook N. *Analysing Musical Multimedia*. Oxford: Clarendon Press, 1998. 304 p. https://doi.org/10.1093/oso/9780198165897.001.0001
10. *Multimedia: From Wagner to Virtual Reality*. Comp. and ed. by R. Packer, K. Jordan; foreword by W. Gibson. New York: W.W. Norton & Company, 2001. 394 p.

11. Kay A. User Interface: A Personal View. In: *Multimedia: From Wagner to Virtual Reality*. Ed. by R. Packer, K. Jordan. New York; London: W.W. Norton & Company, 2001, pp. 121–131.

12. Ascott R. Is There Love in the Telematic Embrace? In: *Multimedia: From Wagner to Virtual Reality*. Ed. by R. Packer, K. Jordan. New York; London: W.W. Norton & Company, 2001, pp. 305–316.

13. Harris L. *Composing Audiovisually: Perspectives on Audiovisual Practices and Relationships*. London: Routledge; New York: Focal Press, 2021. 228 p.

*Information about the author:*

**Pavel A. Michkov** — Cand.Sci. (Arts), Vice-Rector for Research and Digital Development, Associate Professor at the Department of Music Theory and Composition, M.I. Glinka Novosibirsk State Conservatory, Novosibirsk, Russian Federation.

*Информация об авторе*:

**П.А. Мичков** — кандидат искусствоведения, проректор по научной работе и цифровому развитию, доцент кафедры теории музыки и композиции, Новосибирская государственная консерватории имени М.И. Глинки, г. Новосибирск, Российская Федерация.

Received / Поступила в редакцию: 03.08.2026
Revised / Одобрена после рецензирования: 24.08.2026
Accepted / Принята к публикации: 28.08.2026