



<https://doi.org/10.48417/technolang.2024.04.04>

Research article

## Translating Sounds into Visual Images, and Vice Versa

Nina Sokolova  

Peter the Great St.Petersburg Polytechnic University, Polytechnicheskaya, 29, St.Petersburg, 195251,  
Russia

[alonsa@inbox.ru](mailto:alonsa@inbox.ru)

### Abstract

This article outlines the history of theoretical developments, techniques and technologies for translating audio track sounds into graphic elements or pictorial images, and vice versa. The presentation includes various attempts by philosophers (Pythagoras, Aristotle), artists (Guiseppe Arcimboldo, Olivier Messiaen, Alexander Scriabin, Arnold Schoenberg, Wassily Kandinsky, Valentin Afanasyev, Richard David James), natural scientists (Isaac Newton, Robert Hooke, Ernst Chladni, Hans Jenny), engineers (Louis Bertrand Castel, Evgeny Sholpo, Arseny Avraamov, Boris Yankovsky) to reveal the unity of color and sound and to translate colors as well as drawings into sound forms and vice versa. It features a whole variety of such translation types throughout the history of humankind, including mathematical, neuropsychological, physical, technical, and software translation. In the first type, translation is carried out by correlating the wavelength of a particular sound or musical tonality with a certain hue. Its adherents created a color and sound matching table based on their calculated mathematical formulas. Some of them then construct technical devices (like color-harpsichords) that allow this translation to be clearly demonstrated to the public. Within the framework of the neuropsychological approach, the phenomenon of synesthesia is being analyzed as well as the manifestations of this ability for a holistic perception of reality by artists and musicians who tried to convey their experience through works of art in which color and sound are constantly converted into each other. Representatives of physical translation focus on translating sounds into graphic forms and conduct experiments on the effect of sounds of different frequencies and amplitudes on physical substances such as sand, special powder or even liquids, classifying the resulting graphic forms. Adeptes of technical translation are mainly engaged in sound recording, which is carried out by applying certain graphic patterns to film. A modern software approach allows for the translation of sound and image into each other in both directions using digital technologies.

**Keywords:** Sound-Color Translation; Sound of Image; Seeing Sound; Multimodal Perception; Audio-Visual Unity

**Acknowledgment** The author would like to express deep gratitude to Daria Bylieva for her invaluable help and support in writing the article.

**Citation:** Sokolova, N. (2024). Translating Sounds into Visual Images, and Vice Versa. *Technology and Language*, 5(4), 38-58. <https://doi.org/10.48417/technolang.2024.04.04>



© Sokolova, N. This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 1:001(075.8)

<https://doi.org/10.48417/technolang.2024.04.04>

Научная статья

## Перевод из звука в изображение и обратно

Нина Александровна Соколова  Санкт-Петербургский политехнический университет Петра Великого (СПбПУ), Политехническая,  
29, Санкт-Петербург, 195251, Россия[alonsa@inbox.ru](mailto:alonsa@inbox.ru)

### Аннотация

Статья посвящена исследованию истории теоретических разработок, техник и технологий перевода звуков аудиотреков в элементы графики и/или живописные образы, и наоборот. Анализу подвергаются разнообразные попытки философов (Пифагор, Аристотель), артистов (Дж. Арчимбольдо, О. Мессиян, А. Скрябин, А. Шёнберг, В. Кандинский, В. Афанасьев, Р.Д. Джеймс), естествоиспытателей (И. Ньютон, Р. Гук, Э. Хладни, Х. Дженни), инженеров-экспериментаторов (Л.Б. Кастель, Е. Шолпо, А. Авраамов, Б. Янковский) выявить единство цвето-звуковых и/или звуко-графических образов, осуществить перевод цвета/узора в звуковые формы и обратно. Исследуется всё многообразие типов подобных переводов в течение истории человечества, включая такие, как математический, нейропсихологический, физический, технический, программный. В первом типе перевод осуществляется посредством соотнесения длины волны того или иного звука или музыкальной тональности с определенным оттенком цвета. Его адепты на основании высчитанных ими математических формул создают таблицы соответствий цвета и звука. Некоторые из них в дальнейшем конструируют технические приспособления (типа цветных клавишных), позволяющие наглядно демонстрировать данный перевод. В рамках нейропсихологического подхода анализу подвергается феномен синестезии и проявления данной способности к целостному восприятию действительности художниками и музыкантами, пытавшимися передать свой опыт посредством художественных произведений, в которых цвет и звук постоянно конвертируются друг в друга. Представители физического перевода концентрируют внимание на переводе звуков в графические формы и ставят эксперименты по воздействию звуков разной частоты и амплитуды на физические субстанции типа песка, особого порошка или даже жидкости, классифицируя получившиеся графические формы. Адепты технического перевода занимаются, в основном звукозаписью, осуществляя ее путем нанесения на пленку определенных графических узоров. Современный программный подход позволяет осуществлять перевод звука и изображения друг в друга в обоих направлениях с помощью специального программного обеспечения.

**Ключевые слова:** Аудиовизуальный перевод; Окрашенный звук; Звучащие краски; Мультиmodalное восприятие; Синестезия; Единство аудиовизуальных форм

**Благодарность:** Автор выражает глубокую благодарность Дарье Быльевой за неоценимую помощь и поддержку в написании статьи.

**Для цитирования:** Sokolova, N. Translating Sounds into Visual Images, and Vice Versa // Technology and Language. 2024. № 5(4). P. 38-58. <https://doi.org/10.48417/technolang.2024.04.04>



© Соколова Н. This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



## INTRODUCTION

The sound and visual “dimensions” of our world at first glance seem to afford different ways of experiencing, knowing, or revealing reality. They are absolutely dissimilar modalities of perception. The human ability to create auditory sign systems, it would seem, does not intersect with the visual dimension, and there is no possibility of “translating” music into visual form.

However, this is not entirely true. An object can be characterized by visual and auditory features, which in a sense allows one to construct a holistic impression. Throughout human history, many people, including scientists, engineers, composers and visual artists, have sought to convey this holistic image, to reveal the interconnection and unity of audiovisual channels of perception, to provide the ability to translate one modality into another. They did this with different goals: to demonstrate the unity of reality; to enable deaf people to perceive music (Castel) to enrich the methods of artistic influence on the emotional world of recipients (Scriabin, Messiaen, Kandinsky, James); to support with evidence the possibility of influencing the body through various sounds (Jenny); to create sound recording machines (Sholpo, Avraamov); to expand the sound palette of modern music limited by the sounds of existing musical instruments (Yankovsky), etc.

This paper aims to explore the various methods of such “translation” of the progressions of sound into the visual and back, from the first theoretical attempts of antiquity to contemporary methods of conversion that are performed with the help of computer technologies.

### FROM SOUND TO COLOR: MATHEMATICAL TRANSLATION

Since ancient times, people have tried to comprehend and describe the unity of reality by drawing parallels between the visible and audible worlds. The pioneer here was Pythagoras. He suggested a direct connection between audible combinations of sounds and the harmony perceived by the visual organs. According to the ancient Greek thinker, the movement of planets attached to the celestial spheres produces music of extraordinary beauty, which is a mere reflection of the perfection and harmony of the universe. Sounds and visual images are manifestations of the same principle of the unity of the world, which can be described mathematically.

After Pythagoras and his disciples, Aristotle contemplated the relationship between sound and color, and in his *Metaphysics* he claimed that color is measurable by number, and consonance is also a number, a ratio, and the transition through the smallest intervals from the outermost string of the lyre to the highest through intermediate tones is similar to the movement from white – through scarlet and gray – to black (Aristotle, 350 B.C.E / 2013, p. 171).

Therefore, ancient Greek philosophers attempted to understand the essence of the unity of the universe, and one result was that they saw a match between sound sequences and color spectra. However, the value of finding a unified theoretical perspective distracted them from the possibility of achieving unity in a practical way.



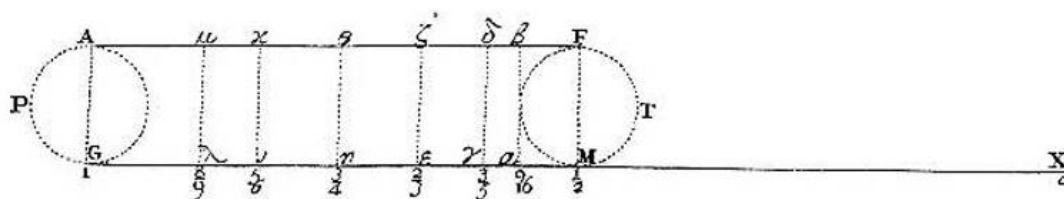
During the Renaissance, the painter Guiseppe Arcimboldo (1527–1593) regained interest in the relationship between sound and color. He began by grading colors from white to black (with a special range of gray gradations) and relating them to the system of harmonic proportions of tones and semitones developed within the Pythagorean school. Arcimboldo laid out a scale of transitions from one shade to another with thinly applied layers of black glazes on a white base and wrote out the corresponding steps of the musical scale. Light color tones corresponded to lower notes, dark ones to higher ones.

Arcimboldo also invented the first ever color clavichord, with each note corresponding to a color. The lowest notes were represented by white, the middle notes by yellow, green, and blue, and the high notes by purple, violet, and bright red.

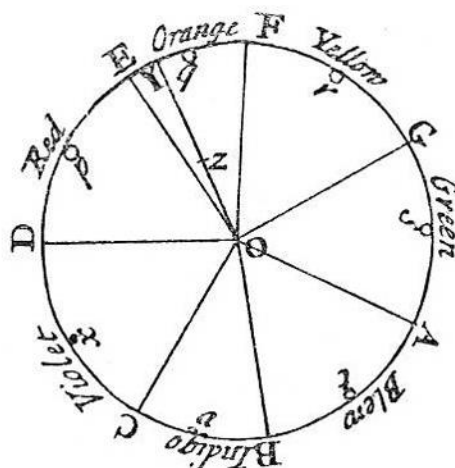
A few years later, the German scholar monk Athanasius Kircher created tables of the relationship between musical notes, colors, their brightness and saturation (*Ars magna lucis et umbrae*, 1646). Later, in his work *Musurgia Universalis* (1650), Kircher developed correspondences between color and musical intervals.

At the end of the 17th century, Newton decomposed sunlight into five colors (red, yellow, green, blue, purple) and proposed a hypothesis that the hue of each color corresponds to a certain wavelength. Assuming that the properties of color and sound are the same, he matched between basic notes and colors, found that two colors were missing and supplemented the color palette of the rainbow with orange and indigo.

Newton wrote: “The Rectilinear Sides MG and FA were by the said cross Lines divided after the manner of a Musical Chord. Let GM be produced to X, that MX may be equal to GM, and conceive GX,  $\lambda X$ ,  $\iota X$ ,  $\eta X$ ,  $\epsilon X$ ,  $\gamma X$ ,  $\alpha X$ , MX, to be in proportion to one another, as the Numbers, 1,  $\frac{8}{9}$ ,  $\frac{5}{6}$ ,  $\frac{3}{4}$ ,  $\frac{2}{3}$ ,  $\frac{3}{5}$ ,  $\frac{9}{16}$ ,  $\frac{1}{2}$ , and so to represent the Chords of the Key, and of a Tone, a third Minor, a fourth, a fifth, a sixth Major, a seventh and an eighth above that Key: And the Intervals  $M\alpha$ ,  $\alpha\gamma$ ,  $\gamma\epsilon$ ,  $\epsilon\eta$ ,  $\eta\iota$ ,  $\iota\lambda$ , and  $\lambda G$ , will be the Spaces which the several Colors (red, orange, yellow, green, blue, indigo, violet) take up” (Newton, 1730, p. 127) (fig. 1, 2).



**Figure 1.** Musical intervals by Newton (Newton, 1730, p. 128)



**Figure 2.** Correspondence between musical intervals and the primary colors of the spectrum, according to Newton (Newton, 1730, p. 155)

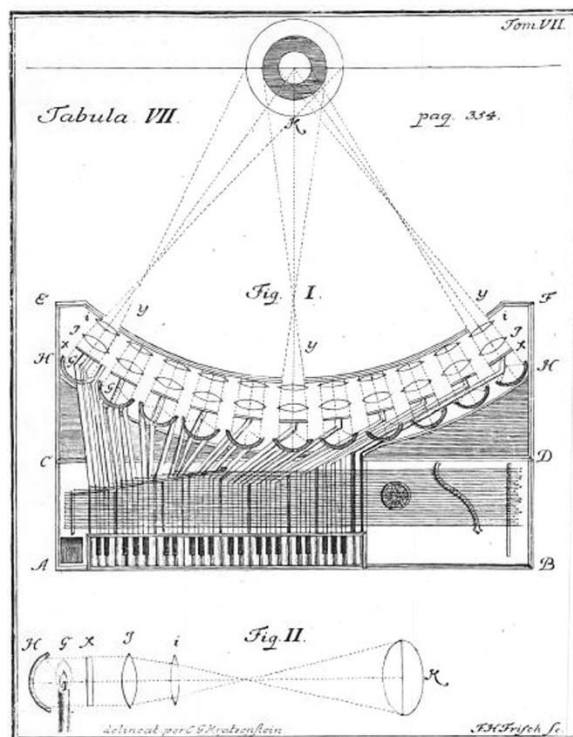
Based on Newton's theory, the Jesuit monk Lois Bertrand Castel proposed in 1734 to create music for the eyes that could be perceived by deaf people. As a result, he invented the color harpsichord (*Clavecin Oculaire*). When it was played it produced no sound but only changes in the color range (Franssen, 1991, p. 33).

The principle of the instrument was as follows: when the musician pressed a key, a panel (or in another version, a crystal illuminated by a candle flame) of the color corresponding to the sound appeared in the frame above the harpsichord. Longer optical waves corresponded to lower frequency sounds. At first, Castel used only seven colors, later he included a range of musical semitones and obtained the following table of correspondences: C – blue, C# – pale green, D – green, D# – yellowish-green, E – yellow, F – peach, F# – orange, G – red, G# – magenta, A – violet, A# – bluish-violet, B – violet-blue.

Later Castel's idea was criticized by Johann Gottlob Krüger, who claimed that it was too subjective to correspond the colors and musical notes in such a way as it is described above. He tended to achieve real harmony in visual representation of music.

Krüger arranged some candles in a shape of semicircle. Each of the candles was placed in the focus of a hollow mirror. “The beams of light coming from the candles were each focused by a lens, such that all the beams projected into one point, the middle of the full circle, where a screen was set up. Each key of the instrument was not only triggering an ordinary harpsichord mechanism but was attached as well to a lever that normally screened off one of the beams, but when moved by pressing the key, pushed a circular window of colored glass into the beam, resulting in the projection of a colored circle onto the screen. The diameters of the windows decreased as the corresponding tones got lower, enabling the simultaneous projection of different colored circles to visualize a colour chord, showing the root of the chord as a primary colour along the circumference of the projected circle and an array of increasingly superimposed colors towards the centre” (Franssen, 1991, p. 38) (fig.3).





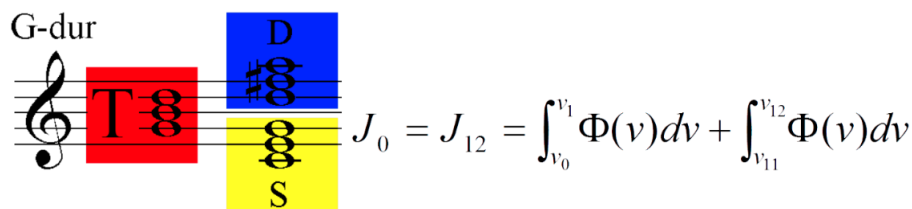
**Figure 3.** Clavecin Oculaire by Johann Gottlob Krüger (Krüger, 1743, plate 7)

In the 20th century, many versions of color organs were created. Among their creators were Alexander Wallace Rimington, Bainbridge Bishop, Mary Hallock-Greenewalt, and others.

In Russia, a representative of the Leningrad underground, Valentin Afanasyev, developed his own system of translating sound into color (1997). Having analyzed previous attempts, he concluded that until now the creators of such translations relied either on their own subjective associations of a particular note with color, or on intuition. From Afanasyev's point of view, the system of sounds (acoustic waves) can be unambiguously correlated with the system of colors (electromagnetic waves) due to the existence of general laws of musical and pictorial harmony.

Both music and painting have a psychophysical effect on a perceiver. Both sound and color are irritants that cause certain tensions in the human psyche. At the same time, tensions of the same intensity caused by the impact of sound or color can be interchangeable. Just as a musical chord can cause certain tensions in us, exactly the same effect is achieved by the impact of color combinations on the eyes.

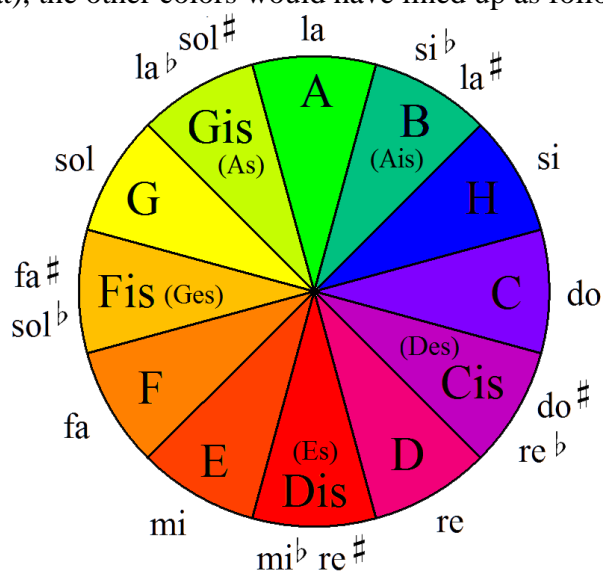
If in music a composer deals with tonal, subdominant or dominant relations (functional), then in the visual arts the basis of relations is the principle of complementarity of red, blue and yellow colors. Afanasyev mathematically calculated and substantiated the mutual proportionality of such relations of music and painting. The following figure shows one of the formulas for his system (fig.4).



**Figure 4.** Common laws of sound and color (Afanasyev, 2024)

According to Afanasyev, certain colors cannot correspond to individual notes. “It is necessary to link the relationships of sounds and colors depending on the plane in which they are presented: melody, harmony or tonality. In other words, the same note in a work can be colored in different colors. Thus, the artist-composer chooses the color tonality of his work, and all subsequent development occurs according to the established laws of sound and color” (Afanasyev, 2024).

The tone color specified by a composer is displayed as a background or the staff on which the notes will be written later. If we take the *Prelude in Es minor* from Volume I of J.S. Bach's *Well-Tempered Clavier* as an example, then, if the author had marked the tonality in red (E-flat), the other colors would have lined up as follows (fig. 5):



**Figure 5.** Correspondence between colors and musical tones (Afanasyev, 2024)

If we take the initial fragment of the *Prelude in Es minor*, we can see a transition from a major to a minor key. This transition could be demonstrated by changing color from warm (red and orange) into cold (reddish purple) (fig.6).



**Figure 6.** Fragment from the *Prelude in E minor* from Volume I of J.S. Bach's *Well-Tempered Clavier* (Afanasiev, 2024)

## NEUROPSYCHOLOGICAL TRANSLATION: AUDIO-VISUAL UNITY

There are people who perceive each hue as correlated with a sound of a certain pitch. “On a closer examination of the multimodal perception, it can be seen that sounds involuntarily trigger colors, lines and shapes in certain human subjects, unified into unique audio-visual experiences” (Borza, 2016, p. 61). Such people are called synesthetes, and their special capacity is chromesthesia, or sound-to-color synesthesia. Whitelaw (2008) called for considering synesthesia not only a neurophysiological, but also a cultural phenomenon, since many synesthetes are creators of synthetic works of art that allow other people who do not have such a rare capability to perceive the world as a whole, in the unity of its audiovisual forms.

Among musicians, this phenomenon was observed in Scriabin, Nikolai Rimsky-Korsakov, Mikalojus Čiurlionis, Messiaen, Schoenberg, Duke Ellington, Billy Joel, Thomas Köner, Tori Amos, and others. Each of them used their synesthetic abilities in their own way.

Synesthetic experience can manifest itself in the use of color references in the titles of musical works, as in Messiaen's *The Colors of the Celestial City*, *Blackbird*, *Chronochromie*, *The Sadness of the Big White Sky*. Or as color references in the score: *The Reed Warbler*, *Quartet for the End of Time* (Johnson, 2023, p. 15).

Synesthesia can also be expressed in a composer's desire to reproduce his experience of the holistic experience of music and color, as was the case with Alexander Scriabin. Scriabin created his own system of color-tonality relationships: “The color underlines the tonality; it makes the tonality more evident” (Myers, 1911, p. 228), where C major was associated with red, D major with yellow, G major with orange-pink, and A major with green.





In 1911, Scriabin wrote his first synesthetic work, *Prometheus: The Poem of Fire*. His idea was to create a work that would unite sound and color. Being a chromesthete, he could not imagine music that would lack simultaneous color expression. The score of *Prometheus* was supplemented with a special performance line, Luce (from Italian for “light”) (fig. 7), which prescribed the sequence of color flashes of light that corresponded to the tonalities being played. It was written in regular notes without any explanation of the correspondence between notes and colors and was to be played using the *Tastiera per luce* (“keyboard with lights”).

## Prométhée.



**Figure 7.** Score of *Prometheus* with the part of Luce (Scriabin, 1911)

In order to bring his idea to life and to demonstrate the idea of color music, Scriabin needed to develop specific equipment. The composer's friend, the engineer Alexander Moser, began working on its creation. His idea was to connect the keys to special bulbs, each of which would be responsible for a particular sound (fig. 8). But this device was never brought to the level necessary for the adequate implementation of Scriabin's project.



**Figure 8.** Moser's color light instrument project (from the collection of Memorial Museum of Aleksander Scriabin in Moscow, Russia)

Enthusiasts from the USA, inspired by the idea of Scriabin's color music, came up with a new way to implement it. For the first time, *The Poem of Fire* with light



accompaniment was performed at Carnegie Hall in New York in 1915, about a month before the composer's death. A gauze screen was stretched over the heads of the musicians onto which multi-colored lights were projected, replacing each other in time with the music.

In Russia, *Prometheus* was first performed in accordance with the composer's original plan only in 1962: streams of light were projected directly onto the walls of the hall, becoming brighter or dimmer depending on the pitch of the notes being played, changing the color palette depending on the tonality.

Another composer and sometimes painter, Arnold Schoenberg (1874–1951), also conceived of the mutual enrichment of sound and color in a *Gesamtkunstwerk* (“total work of art”). If we try to “translate details of musical language, which reason cannot understand, into the language of our concepts the essence is lost” (Kandinsky & Marc, 1912/2005, p. 92). To convey this essence, to present it differently but still coherently, can only be done by using the language of another art.

In his monodrama *The Happy Hand* (1924), Schoenberg created an illuminated-scene design, which turns out to be subordinate to the music. In the musical score, he clearly wrote down his instructions regarding the stage design, costumes, appearance of the characters, as well as the nature and color of the lighting. The composer personally made sketches for the drama, conveying his color ideas, describing in detail the moments of changing lighting during this or that action of the characters on stage.

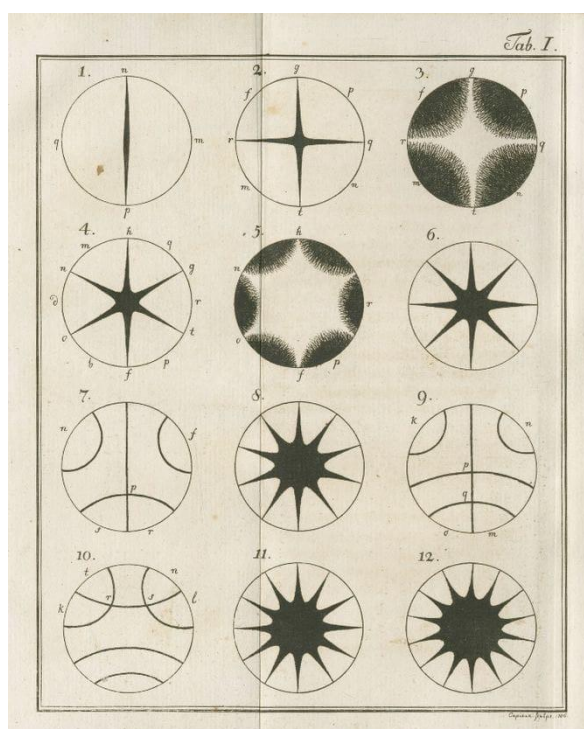
Some artists also sought to bridge the gap between sounds and colors by hearing how colors “sound.” Thus, Wassily Kandinsky used his gift for hearing colors in his work to create paintings such as *Musical Overture* (1919) *The Violet Wedge* (1919), *Composition IV* in his essay on Scriabin, and *Composition VIII* (1923), inspired by Mahler, as well as a tableau entitled *Parallels of Color and Sound* (Rucsanda, 2019). He also created four “color-tone dramas”: *The Yellow Sound*, *The Green Sound*, *Black and White*, and *Violet*. The opera *Der Gelbe Klang (The Yellow Sound)* (1912) had no plot in the usual sense, but was a mixture of color, light, and sound, featuring five “bright yellow giants (as big as possible)” and “vague red creatures, somewhat reminiscent of birds” (Casini, 2017). The action was described in it through constant cross-references between the sounds of the orchestra and the colored beams of spotlights illuminating the figures on stage: “Suddenly all colors vanish (the giants remain yellow), and a dim white light fills the stage. In the orchestra single colors begin to speak. Corresponding to each color sound, single figures rise from different places” (Kandinsky & Marc, 1912/2005, p. 223).

In his text *On the Spiritual in Art*, Kandinsky draws parallels between colors and musical instruments: “Blue, presented musically, resembles a flute, dark blue resembles a cello and, becoming darker, the wonderful sounds of a double bass; in a deep, solemn form, the sound of blue can be compared to the low notes of an organ” (Kandinsky, 1967, p. 96). White “sounds like a non-sound, which quite accurately corresponds to certain pauses in music <...>. Presented musically, black is a complete final pause” (p. 101). “The light, warm red color resembles the sound of a fanfare with a tuba overtone <...>. Red cinnabar sounds like a tuba” (p. 104). Deep green – like a cello. Cold red – like a violin. Orange – like an alto violin. Violet – like an English horn, bassoon.



## PHYSICAL TRANSLATION: NATURE-DRAWN SOUND

In the 17th century, scientists noticed the ability of sounds to imprint themselves in the form of patterns. Robert Hooke was one of the first to draw attention to the patterns on loose surfaces from sounds. In 1680, while running a bow over a metal plate, he discovered patterns that were formed from the flour scattered on it. Ernst Chladni expanded his experiments in 1787, creating an entire encyclopedia of patterns that different sound waves leave on sand (Faraday, 1831) (fig. 9).



**Figure 9.** Chladni's figures (Chladni, 1787, p. 87)

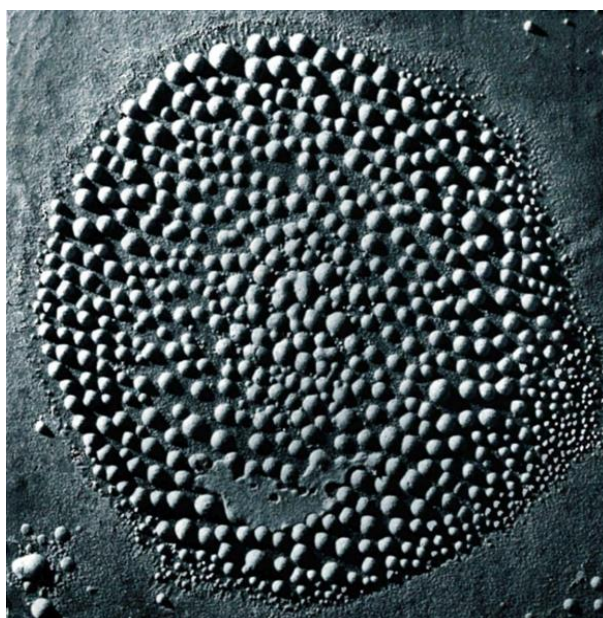
Inspired by Chladni's ideas, Swiss physician Hans Jenny created the science of cymatics (from the Greek κύμα – “wave”) in 1967. His aim was to study the effect of sounds on the human body. Having improved Chladni's technique, he invented a device called a tonoscope, capable of converting any sound (including the human voice) into an image. Jenny recorded individual vowels in different languages of the world using a tonoscope and noticed that when they were pronounced in ancient languages (for example, Sanskrit), the sand took the form of written symbols corresponding to the vowels being pronounced.

Using lycopodium powder, Jenny also created moving 3-D objects (fig. 10, 11).





**Figure 10.** Loud sound (large amplitude), lifting the powder up (Jenny, 2001, p. 69)



**Figure 11.** Acoustic irradiation transforms a layer of lycopodium powder into round shapes every one of each rotates on its axis and around the whole figure (Jenny, 2001, p. 75)

### TECHNICAL TRANSLATION: GRAPHICS TO SOUND

Sound recording technologies showed the connection between sound and the visual in a new way. Sound received its own specific visual form, not yet connected with color. In 1889, in Russia, Vekshimsky invented a device for optical sound recording. The vibrations of the membrane caused by the sound are transmitted to a mirror, which reflects a light beam. The beam, in turn, passes through a thin slit, forming a “light stroke,” and the set of strokes of different heights makes up a picture of sounds as mountain peaks. In 1904, Eugene Augustine Last presented a prototype of a system for optical sound recording on film, recording the definiteness of sound as changing waves.



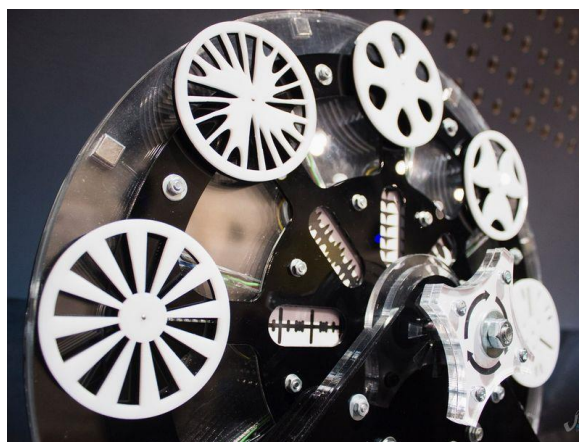
In 1919, American inventor Lee De Forest received a patent for a film sound recording process in which he improved the design of Finnish inventor Eric Tigerstedt and the German Tri-Ergon system and called this process *Forest's Phonofilm*. In *Phonofilm*, sound is recorded directly onto the film in the form of a track of variable optical density, in contrast to the “variable width” method in the *Photophone* system developed by RCA. Changes in the track density correspond to a pulsating current of audio frequency from a microphone and are applied photographically to the film, and during the screening of the film are converted back into an electrical signal by a photocell. Thus, the idea of converting an audio signal into a visual one and back received its further development, as well as practical application.

At the same time, Soviet engineers were conducting their own research to convert sound into graphic form. As a result, two systems with a low-inertia galvanometer were developed in Moscow and Leningrad at almost the same time: with a variable width of the optical track by Alexander Shorin, and with a variable density phonogram *Tagefon* created under the supervision of Pavel Tager.

The overall scheme of recording is as follows: the sound of the amplified microphone is transmitted to special equipment connected to the incandescent lamp. The rotation (corresponding to rhythm and tone) of the light emitted from it passes through a special lens and is recorded on the moving photosensitive film.

Studying the visual imprint of sound, Soviet researchers in the 1920s began to think about “reverse translation.” If there is a visual representation that uniquely corresponds to a voice or audio recording, then could decorative patterns be “translated” into music? There were several researchers in the Soviet Union working in this direction.

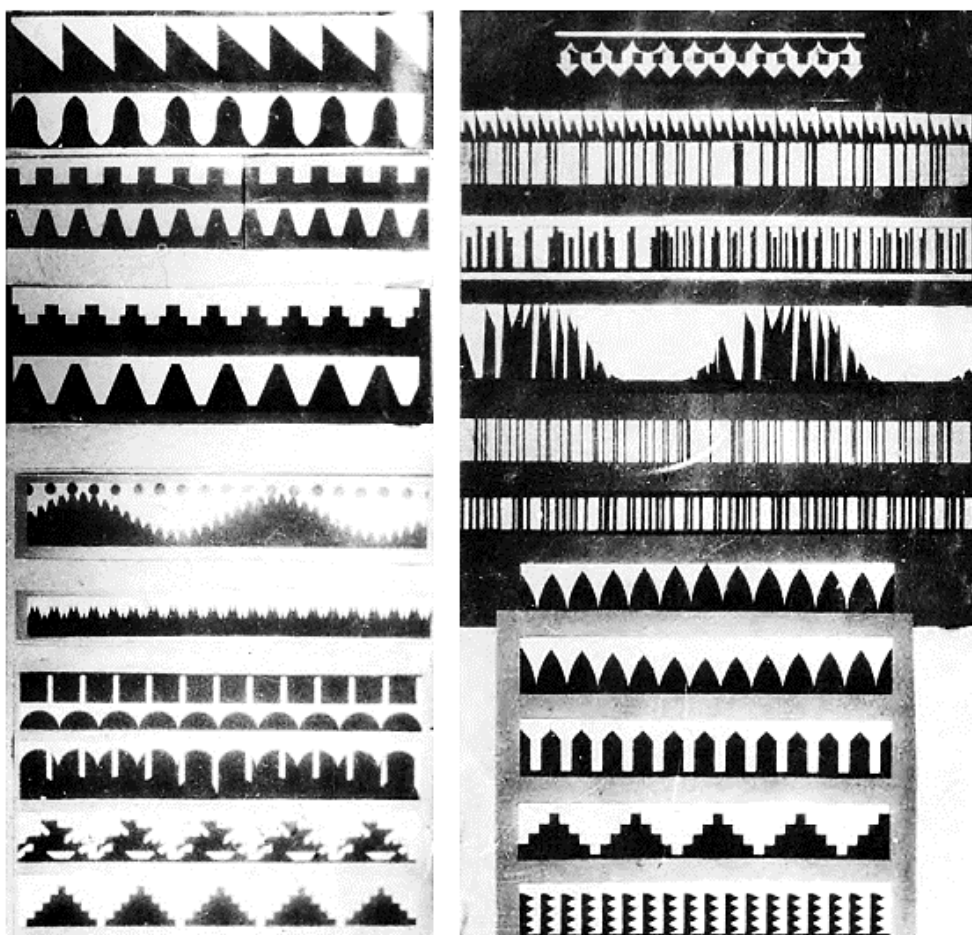
In 1931, Sholpo designed the *Variophone*. It consisted of cardboard disks attached to a rotating circle with teeth of various shapes cut into them, forming one or another sound wave. As they rotated, these disks periodically interrupted a beam of light, which formed the outlines of the soundtrack. A beam of light emanating from a projector shone through the rotation of these disks and was recorded on film as a picture. Then these images, passed through a sound projector, were converted into sound (fig. 12).



**Figure 12.** “Variophone” by Sholpo with cardboard discs with basic wave shapes (from the collection of the Museum of Sound in St. Petersburg, Russia)



If Sholpo, before creating his device, graphically recorded and combined various sounds and consonances, composer Arseny Avraamov immediately began with the use of ornaments. He drew them on paper (fig. 13), then photographed them on the soundtrack of a film and reproduced them using a projector. Moreover, if the melody of the resulting compositions was recognizable, then their timbre was unique. It seemed as if the music was voiced by some unknown instruments.



**Figure 13.** Patterns for translation graphics into sounds (see Smirnov, 2013, p. 179)

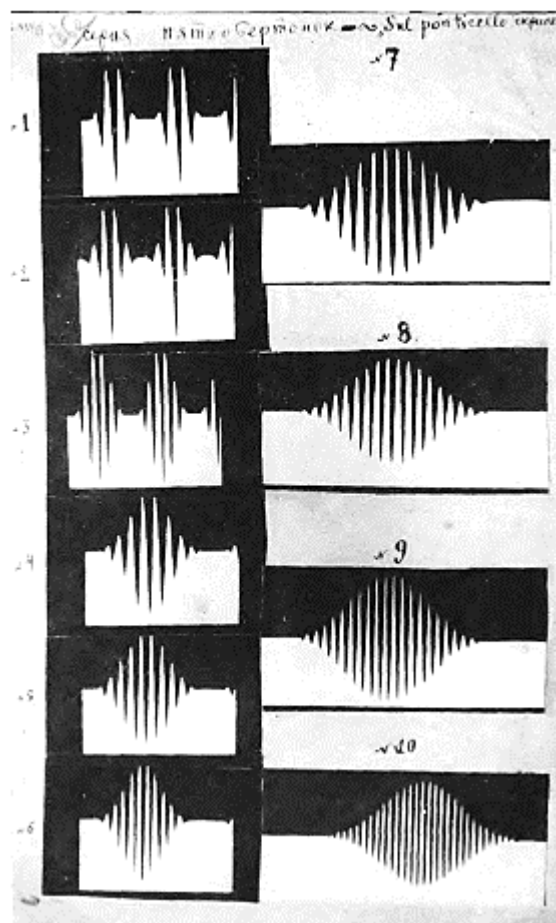
Later, Avraamov tried to translate into sound more complex geometric figures such as algebraic equations and images of molecular movements within certain chemical elements.

Unlike his colleagues, Boris Yankovsky wanted to achieve the definition of not only the form of the sound wave (which was dictated by the ornament), but also the timbre. His plans included the creation of new tonal systems, complex polyrhythmic effects. For this purpose, he proposed to assemble a collection of sound elements similar to the periodic table (fig. 14) and on its basis form a universal language of sound which would be studied by a new discipline – *Synthetic Acoustics*. According to Yankovsky, just as the gaps in the periodic table are gradually filled thanks to the latest discoveries in





chemistry, with the help of *Synthetic Acoustics* it is possible to fill the gaps in the system of orchestral tonal colors. He proposed to do this by selecting and crossing the sounds he selected and recording sound waves using sinusoids.



**Figure 14.** Sounds of synthetic instruments that were supposed to fill the gaps between the sections of a symphony orchestra (see Smirnov, 2013, p. 179)

Animator Nikolai Voinov cut out sound wave profiles from paper, synthesized them optically using the *Nivoton* tool he created, and then photographed fragments of the soundtrack on an animation machine, converting the image into sound. By combining them with the video sequence, he received a full-fledged synthetically voiced animated film.

### **SOFTWARE TRANSLATION: FROM SOUND TO IMAGE, AND BACK**

Digitalization has significantly expanded the possibilities for “translating” audio and visual forms into each other. In the 20th and 21st centuries, special programs for converting sound into graphics/color and back are created using computer technology. The principle of operation of such programs is as follows. Every sound has two main

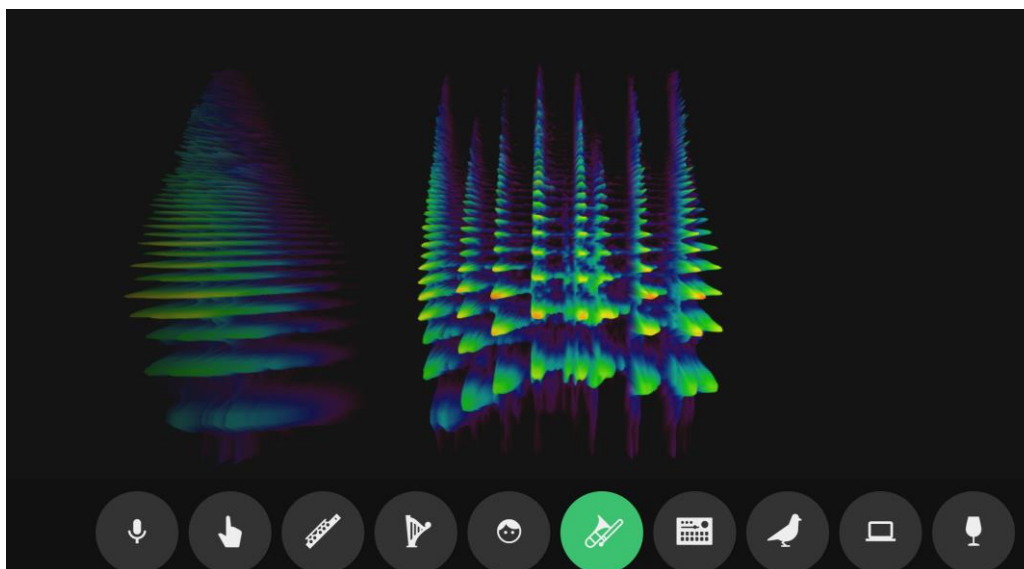


parameters: frequency (which determines the pitch of sound) and amplitude (responsible for the volume). In a spectrogram (a visual display of sound), frequencies are displayed on the vertical axis (Y), amplitude is designated by one color or another, and the horizontal axis (X) reflects the time characteristics of the audio track.

At present, such computer programs use different variants of sound-to-visual image conversion (their differences lie in different degrees of scaling, compression and stretching of axes). They are used for medical purposes (ultrasonic examinations, which result in a visual image that is easier for a doctor to comprehend and interpret than audio recording), in the analysis of various audio data obtained from nature (for example, birdsong or roaring of some wild animal). These programs could be also used for automatic addition of sound effects to videos for the purpose of reducing manual sound editing work (Zhou et al., 2018).

At present, such computer programs use different variants of sound-to-visual image conversion (their differences lie in different degrees of scaling, compression and stretching of axes). They are used for medical purposes (ultrasound examinations, which result in a visual image that is easier for a doctor to comprehend and interpret), in the analysis of various audio data obtained from nature (for example, birdsong or roaring of some wild animal).

There are browser programs that allow a person to see various sounds in color and three-dimensional representation (for example, fig. 15, to “draw music”, that is, by adjusting the key, octave, etc., one can then move the cursor across the screen, which will be displayed both visually and audibly (for example, <https://spectrogram.sciencemusic.org/>)

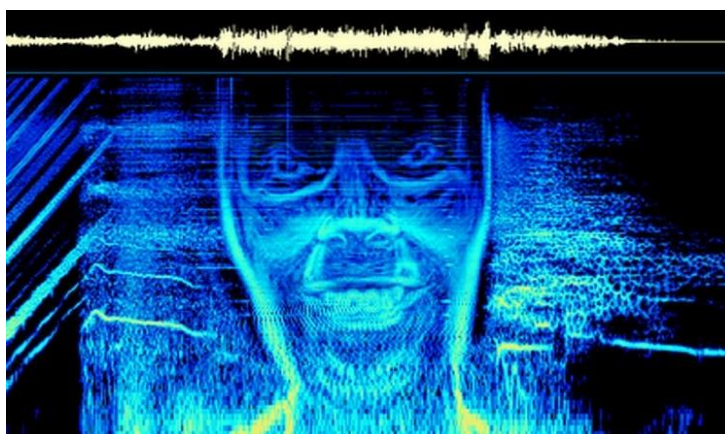


**Fig. 15.** Visual representation of saxophone sounds (Jeremy Morril and Boris Smus, <https://musiclab.chromeexperiments.com/spectrogram-service/>)



The *Increat* program (2003) allows translation between audio and visual data based on the analogy between the main parameters of sound and color. Thus, red shades of the image correlate with the left channel of the stereo sound system, green – with the right, and yellow – with both channels. Such sound characteristics as pitch, volume, duration of sound correspond to brightness, color intensity, a certain length and shape of the line.

Some musicians have begun to use similar technologies in their works. For example, Richard David James (*Aphex Twin* art project), having created images, records them as a melody, then mixes these “pictures” of audio files with other sounds, creating unique tracks. By passing these music tracks through a spectrograph, it is possible to perform the reverse operation of obtaining a picture from sound. Thus, by inserting a spectrogram into the soundtrack of his composition *Equation*, James captured the “face of a demon”. As it turned out later, this image was a modified self-portrait of the musician. By changing some settings in Spectrogram, a program that allows visualization of the original soundtrack, user Jarmo Niinisalo saw and showed to other people the face of the creator of the composition himself (fig. 16).



**Figure 16.** Original “Face of the Deamon” in the soundtrack *Equatio* by Aphex Twin, obtained using a linear frequency display scale instead of a logarithmic one (Wadsworth, 2016)

Thus, new technologies make it possible to create multi-layered meanings in a work of art, accessible to the perception only of advanced users, stimulating them to actively participate in the creation, to make efforts to decipher and translate auditory information into visual information, hinting at the presence of deeper meanings hidden behind the original melody or image.

Another software called Virtual ANS is a graphic editor simulating the photoelectronic synthesizer ANS which was created by Evgeny Murzin in 1958, named after Alexander Nikolayevich Skryabin. It allowed for the first time to draw music in the form of a spectrogram without the participation of live instruments and performers. This is a graphic editor that gives the user the ability to turn sounds into images, download and listen to pictures, draw microtonal/spectral works with a unique cosmic sound.



Recently, software developments have continued to convert sound into image and back. At the same time, they are becoming increasingly focused on the average user, who is constantly connected to their smartphone.

Thus, one of the latest developments is the PhonoPaper camera application created on the basis of Virtual ANS, which provides the ability to listen to images with encoded sound (PhonoPaper codes). In the application, you can also create your own codes: a sound no longer than 10 seconds will be recorded from the microphone and converted into an image.

Browser synthesizers allow us to experiment with converting images into sound. The most common and easiest way to convert sounds is to use individual lines or a set of dots of different heights.

Among Google's experiments is "Paint with music" (<https://g.co/arts/BGBT8h2p4QqnnCPW6>). Here a person can select a specific background, and using one or more of the suggested tools, draw lines that will then be played according to the pitch of the sound, visually exploding at the playing point with a multitude of multi-colored circles.

Inspired by the ANS optical synthesizer, an artist Olivia Jack created a simple web-based graphical synthesizer called Pixelsynth. By default, a person can select one of the pictures drawn on a black background or even text (Fig. 17), which is gradually played from left to right (shown by the red stripe). Accordingly, by drawing in white, you can change the sound.



**Fig. 17.** Pixelsynth web synthesizer that plays white-on-black drawings (Olivia Jack, <https://ojack.xyz/PIXELSYNTH/>)

In general, the digitalization of visual and auditory objects has brought their nature closer together, which has expanded the possibilities for their transformation. For example, the tone value of each pixel in a drawing can be converted into a sound



frequency. Artificial intelligence is increasingly becoming an intermediary in translation. In some cases, the AI image-to-music translation is mediated by text, in others the translation is direct (for example, The Synesthetic Variational Autoencoder SynVAE translates images into music in such a way that objects that are similar in color and location sound similar (Müller-Eberstein & Noord, 2019).

## CONCLUSION

Despite the difference between the visual and auditory modes of perception of reality, ideas about their complementarity, as well as the possibility of mutual translation of sound and visual forms, has interested humanity since ancient times.

At first, theoretical attempts at such a translation appeared, with the help of mathematics. Pythagoras, Aristotle, and then Arcimboldo, Newton, Castel and other scientists sought to establish correspondences between individual sounds and the colors of the spectrum based on calculations of the length of the sound wave and the corresponding hues of color. Sometimes such attempts also had practical application in the form of various musical instruments capable of simultaneously reproducing a sound and the color corresponding to it.

Other scientists proposed a different version of such translation – a physical one. As a result of experiments with various substances, they discovered that sounds, acting on these substances, are capable of creating certain patterns on them. The result of such research was the creation of a device (tonoscope) with the help of which any sound can be translated into an image. The purpose of its creation was, first of all, the desire to improve and facilitate the decoding of the results of medical examinations. Nevertheless, representatives of other professions became interested in the possibility of such translation and began to use it for their own purposes.

Over the past three centuries, artists have also been interested in the possibility of translating musical and visual series. This holds especially for artists who experience synesthesia. They developed a neuropsychological mode of translation that is based on the innate ability of some people to perceive sounds as colored, allowing for the establishment of correspondence between the pitch or timbre of a sound and certain shades of color. These “translations” lead to the emergence of works of art that reflect a holistic vision of the world.

There is also a technical variant of audiovisual data translation, connected with the optical system of sound recording on film with the help of special equipment. Its adherents attempted to translate graphic patterns into a soundtrack. The result of which was, first of all, sound recording equipment

In the contemporary world, software translation dominates, converting sound into images and returning them through computer software.

The digital conversion of auditory and visual signals provides a unified language in which both can be expressed. It will be important to more closely consider how this unification related to other approaches to the unification of sight and sound, including the mathematical/theoretical approach of the Pythagorean. Also, it is open to debate whether this unification provides a synesthetic holistic vision of a world that is digital or can be





digitized throughout. If physicians and physicists drew on the inherent similarity of light waves and sound waves and their resonance which would materially affect physical bodies, there is now a move from a common wave nature to a common representational form, namely that of the digital.

## REFERENCES

- Afanasiev, V. (2024). “Tsvetnoy sluh” i intuitivnye poiski [“Colored auditory perception” and intuitive search]. <https://afanasieff.ru/>
- Aristotle (2013). *Metaphysics*. Roman Roads Media. (Original work published 350 B.C.E)
- Casini, S. (2017). Synesthesia, Transformation and Synthesis: Toward a Multi-sensory Pedagogy of the Image. *The Senses and Society*, 12(1), 1–17. <https://doi.org/10.1080/17458927.2017.1268811>
- Chladni, E. F. F. (1787). *Entdeckungen über die Theorie des Klanges* [Discoveries about the Theory of Sound]. Weidmanns Erben und Reich. <http://echo.mpiwg-berlin.mpg.de/MPIWG:EKGK1SP1>
- Faraday, M. (1831). *On a Peculiar Class of Acoustical Figures, and on the Forms of Fluids Vibrating on Elastic Surfaces*. Printed by Richard Taylor.
- Franssen, M. (1991). THE Ocular Harpsichord of Louis-Bertrand Castel. The Science and Aesthetics of an Eighteenth-Century Cause Célèbre. *Tractrix* 3, 15–77.
- Jenny, H. (2001). *Cymatics. A study of Wave Phenomena and Vibration*. MACROmedia.
- Johnson, R. (2023). Music & Synesthesia: An Exploration of Synesthesia and its Relation to Musical Perception. Pacific Undergraduate Research and Creativity Conference (PURCC). <https://scholarlycommons.pacific.edu/purcc/2023/events/49>
- Kandinsky, W. & Marc, F. (Eds.). (2005). *Der Blaue Reiter. Almanac* [The Blue Rider. Almanac.]. MFA Publications. (Original work published 1912) <https://archive.org/details/blauereiteralman0000unse/>
- Krüger, J. G. (1743). De novo musices, quo oculi delectantur, genere [On a New Kind of Music that Delights the Eyes]. In *Miscellanea Berolinensia, ad incrementum scientiarum ex scriptis Societati Regiae Scientiarum exhibitis* (vol. 7, pp. 345-357). Halae-Magdeburgicae.
- Müller-Eberstein, M., & Noord, N. van. (2019). Translating Visual Art into Music (arXiv:1909.01218). arXiv. <https://doi.org/10.48550/arXiv.1909.01218>
- Myers, C. S. (1911). A Case of Synesthesia. *British Journal of Psychology*, 4(2), 228–238. <https://doi.org/10.1111/j.2044-8295.1911.tb00045.x>
- Newton, I. (1730). *Optics: or a Treatise of the Reflections, Refractions, Inflections, and Colors of Light. First Book, Part 2*. William Innys.
- Rucsanda, M. D. (2019). Aspects of the relationship between Music and Painting and their influence on Schoenberg and Kandinsky. *Bulletin of the Transilvania University of Braşov, Series VIII: Performing Arts*, 12(2), 91–100.





- Scriabin, A. (1911). *Le poeme du feu, Prometheus, Op. 60 (orchestra music score)* [The Poems of Fire, Prometheus, Op. 60 (orchestra music score)]. Editions Russes de Musique.
- Smirnov, A. (2013). *Sound in Z. Experiments in Sound and Electronic Music in Early 20<sup>th</sup> Century Russia*. Koenig Books.
- Wadsworth, J. (2016, June 21). The Devil is in the Detail: Video Game Soundtracks and Spectrograms. *The Oxford Culture Review*.  
<https://theoxfordculturereview.com/2016/06/21/the-devil-is-in-the-detail-video-game-soundtracks-and-spectrograms/>
- Whitelaw, M. (2008). Synesthesia and Cross-Modality in Contemporary Audiovisuals. *The Senses and Society*, 3(3), 259–276.  
<https://doi.org/10.2752/174589308X331314>
- Zhou, Y., Wang, Z., Fang, C., Bui, T., & Berg, T. L. (2018). Visual to Sound: Generating Natural Sound for Videos in the Wild. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3550-3558). ACM  
<https://doi.org/10.1109/CVPR.2018.00374>

#### СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

Нина Соколова, [alonsa@inbox.ru](mailto:alonsa@inbox.ru)  
ORCID 0000-0001-8156-1253

Nina Sokolova, [alonsa@inbox.ru](mailto:alonsa@inbox.ru)  
ORCID 0000-0001-8156-1253

Статья поступила 14 сентября 2024  
одобрена после рецензирования 5 декабря 2024  
принята к публикации 19 декабря 2024

Received: 14 September 2024  
Revised: 5 December 2024  
Accepted: 19 December 2024