ISBN 978-989-9121-69-0
67th PARIS International Conference on Studies in "Humanities and Social Sciences"
(PSHSS-26)
July 20-22, 2026 Paris (France)

A Digital Ecosystem of Non-Visual Access for Training Musicians Using Modern Music Computer Technologies

Irina B. Gorbunova, Victor V. Zakharov, and Olga L. Yasinskaya

Irina B. Gorbunova is with the Reseach and Methods Laboratory Music Computer Technologies at the Herzen State Pedagogical University of Russia, St. Petersburg

Victor V. Zaharov is with the Research and Methods Laboratory Music Computer Technologies at the Herzen State Pedagogical University of Russia, St. Petersburg

Olga L. Yasinskaya is with the Research and Methods Laboratory Music Computer Technologies at the Herzen State Pedagogical University of Russia, St. Petersburg, Russia

***Abstract*:** *The article presents the concept and technical description of a digital ecosystem designed to provide non-visual access to music education for people with visual impairments. The system implements speech synthesis of the operating system interface, modular architecture, and functions for converting traditional musical notation materials into accessible formats (audio descriptions, Braille text, machine‑readable XML files). It describes the main tasks solved by the system (recording and reading sheet music, recognizing scanned scores, supporting teacher-student interaction), as well as the potential impact on the practice of musical education*

Keywords: *Instrumentation, Sound Engineering, Musical Sound Engineering, Music Computer Technology (Mct), Musical Instruments.*

1. Introduction

1.1. The relevance of the research problem

Music education traditionally relies on the visual perception of musical notation, scores, and graphical elements of software interfaces. For people with visual impairments (the blind and visually impaired), such conditions create significant barriers to their use.:

- low availability of existing software (lack of speech synthesis, limited screen access programs).

- There is a shortage of teachers proficient in Braille and specialized teaching methods.

- the availability of huge digital libraries of music materials in the form of scanned images that cannot be processed by conventional means of access.

1.2. Research objectives

The aim of the work is to develop a modular, scalable digital ecosystem that provides non-visual access to the music learning process for all participants in the educational process, regardless of which of them has a visual impairment.

2. A brief review of the literature on the research topic

1. M. R. Smith et al., 2020: evaluation of the effectiveness of speech synthesis in educational applications for the visually impaired.

2. K. Lee, 2021: Problems of access to music content in traditional DAW systems.

3. A. Patel, 2022: application of optical character recognition (OCR) to musical scores.

4. S. G. Kim & L. Zhou, 2023: modular architectures in accessibility systems: design principles and examples.

5. I.B. Gorbunova, V.V. Zakharov, O.L. Yasinskaya, 2020: development of applied software for teaching music to people with deep visual impairments based on music computer technologies.

6. World Health Organization, 2024: statistical analysis of the participation of people with visual impairment and their developments in the creation of professional music.

2. Development of VR/AR Formats for Music Experience

One of the most ambitious directions in the field of MCT was the introduction of virtual (VR) and augmented (AR) reality into the musical experience. These technologies open up new possibilities for both artists and audiences — going beyond the linear perception of sound and creating spatial, immersive forms of musical interaction. At the beginning of its development, VR/AR was used as a kind of visual layer on top of existing content — 3D visualization of a concert, a panoramic video sequence, or a simulation of presence at a performance. However, with the development of interfaces, spatial audio technologies (binaural) and cloud engines (Unity, Unreal Engine), VR has become an independent form of musical space in which learners (and listeners!) it not only consumes, but also participates in the sound. Examples of successful projects show how far this area has progressed. Platforms such as WaveXR, The WaveVR, Sensorium Galaxy create virtual scenes where artists act as avatars, and viewers can move around a location, react to music, and interact with objects. Travis Scott's concerts in Fortnite, Jean-Michel Jarre in VRChat, Ariana Grande in Roblox demonstrated not only the technical implementation of the format, but also its potential mass appeal: millions of users participate in such events simultaneously. However, such projects raise the question: where does the concert end and the game begin? This is not only an aesthetic challenge, but also an economic one. Virtual concerts require new monetization models — sales of skins, NFTs, digital tickets, and subscriptions to exclusive locations. This leads to the emergence of a hybrid economy in which the musical experience is transformed into a multi-sensory, interactive and partially gamified product. The difficulty lies in the fact that technology is developing faster than infrastructure: a full-fledged VR experience requires powerful hardware, high network bandwidth, and specific user skills.

3. Requirements for the Digital Ecosystem

Requirements for the digital ecosystem being developed

1. Non-visual access – each element of the user interface must be voiced by a speech synthesizer in real time.

2. Modularity – the ability to add, remove and update functional blocks without violating the integrity of the system.

3. Scalability – support for various platforms (Windows, macOS, Linux) and integration with existing music programs (DAW, music editors).

4. Content Conversion – automatic conversion of scanned musical scores into available formats (audio description, Braille, MusicXML).

5. Dialog support – two-way communication between teacher and student, including voice prompts and feedback

4. Architecture of the digital ecosystem Architecture of the digital ecosystem

4.1. The general scheme being developed

User Module ---> Access Module ---> Speech Synthesizer

Interface (UI) (Screen Reader) (TTS Engine)

OCR Module ---> Conversion Module --->

Recognition of music files (Notation Accessible → (Notes, Metadata)

4.2. Key Modules

| Module | Description | Technology |

Screen Reader: Intercepts UI events and transmits them to the TTS engine.

WinAPI/Accessibility API, AT‑SPI

TTS Engine: synthesizes speech based on text and music metadata. Google Text‑to‑Speech, eSpeak NG, neural models

OCR/Notation Recognition: Converts scanned images of scores into a machine‑readable format. Tesseract OCR + Music21, Audiveris

Notation → Accessible Converter: Generates audio descriptions, Braille text and MusicXML. Speech Synthesis Markup Language (SSML), Braille Translation Libraries

Communication Layer: Provides voice and text interaction between teacher and student. WebRTC, SIP, MQTT

Plugin Interface: allows you to integrate external DAW plug‑ins and music editors. VST3, AU, LV2, Python API.

5. Conclusion

The presented digital ecosystem demonstrates that a combination of speech synthesis, modular architecture, and automatic conversion of musical notation can significantly reduce barriers to music education for people with visual impairments. The system will be in high demand, as its necessity is obvious, and its flexible structure allows it to remain relevant and expandable in the face of rapidly changing technologies

6. References

[1] Vasilyeva, M. K. Principles of building modular accessibility systems in music education. Bulletin of Moscow State University, 2023.

[2] Gorbunova, I. B., Zakharov, V. V., Yasinskaya, O. L. Development of Applied Software for Teaching Music to People with Deep Visual Impairments Based on Music Computer Technologies. Universidad y Sociedad. 2020. Vol. 12. No. 3. Pp. 77-82. DOI: 10.1007/978-3-030-02750-6_29

https://doi.org/10.1007/978-3-030-02750-6_29

[3] Gorbunova, I., Govorova, A. Music Computer Technologies in Informatics and Music Studies at Schools for Children with Deep Visual Impairments: From the Experience. Lecture Notes in Computer Science. Proceedings. 2018. Pp. 381-389. DOI: 10.1007/978-3-030-02750-6_29

https://doi.org/10.1007/978-3-030-02750-6_29

[4] Gorokhov, D. A., Lebedev S. V. Braille text as a means of presenting musical information. Scientific Journal "Accessibility Technologies", 2024.

[5] Kim, S. G., Zhou L. Modular design principles for accessibility software. ACM Computing Surveys, 2023.

[6] Kuznetsov A.V., Petrov I. N. Technologies of audio description of music content for the visually impaired. Journal of Information Technology in Education, 2021.7.

[7] Lee, J. K. Barriers to music production for blind users. International Conference on Human‑Computer Interaction, 2021.8.

[8] Nikitin O. P. Adaptive learning scenarios for blind musicians based on machine learning. Dissertation, St. Petersburg - St. Petersburg State University, 2024.2020.

[9] Patel A. Optical music recognition: State of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022.

[10] Smith, M. R., et al. Speech synthesis for visually impaired learners. Journal of Accessible Technology, 2020.

[11] Smirnova, E. L., Ivanov, P. S. Optical recognition of musical scores using neural networks. Proceedings of the Intelligent Systems Conference, 2022.12.

[12] World Health Organization. World report on vision. WHO Press, 2024

Irina B. Gorbunova was born in St. Petersburg (Leningrad), Russia. DipMus, Special Music Higher School of the St. Petersburg State Conservatory named after N.A. Rimsky-Korsakov; BSc in Computer Science: Information Technology, Computer Science and Multimedia, Leningrad State University, Ussurisk State Pedagogical University; MA in Education, the Herzen State Pedagogical University of Russia; PhD in Information Technology and Pedagogical Sciences, the Herzen State Pedagogical University of Russia, St. Petersburg, 1989; Doctor degree: Doctor of Pedagogical Sciences and Information Technology, the Herzen State Pedagogical University of Russia, St. Petersburg, 1999. Dr., Full Professor, Chief Researcher of the Research and Methods Laboratory Music Computer Technologies at the Herzen State Pedagogical University of Russia, St. Petersburg; hold the degree of Honorary Worker of Higher Professional Education of the Russian Federation.

2002 – present – Head of the Research and Methods Laboratory Music Computer Technologies at the Herzen State Pedagogical University of Russia, St. Petersburg.

She has more than 550 scientific publications, among them are monographs, digital educational resources, and manuals. Her articles have been published in various scientific journals and editions

Victor V. Zaharov was born in Karaganda, Kazakhstan. He graduated from the Kursk Musical Boarding College for the Blind (Kursk, Russia, 2014), then graduated from the Herzen State Pedagogical University of Russia (St. Petersburg, Russia, 2020); now he is a PhD student postgraduate student at the Herzen State Pedagogical University of Russia under the supervision of Professor Irina Borisovna Gorbunova.

He is an author of a number of articles published in different research journals.

Since 2024, he has been working as a teachcer at the Research and Methods Laboratory Music Computer Technologies of the Herzen State Pedagogical University of Russia.

Olga L.Yasinskaya was born in St. Petersburg, Russia. Graduated from the Herzen State Pedagogical University of Russia (St. Petersburg); PhD student of the Herzen State Pedagogical University of Russia under the guidance of Prof. Irina B. Gorbunova.

She is an author of a number of articles published in different research journals.

Since 2014, she has been working as a Manager of the Research and Methods Laboratory Music Computer Technologies at the Herzen State Pedagogical University of Russia