



O'zbekiston Respublikasi  
Vazirlar Mahkamasi  
huzuridagi O'zbek tilini  
rivojlantirish jamg'armasi



O'zbekiston Respublikasi  
Raqamli ta'lim  
texnologiyalar  
vazirligi



O'zbekiston Respublikasi  
Oliy ta'lim, fan va  
innovatsiyalar vazirligi



Sankt-Peterburg davlat  
universiteti



Istanbul texnika  
universiteti



Toshkent davlat o'zbek tili  
va adabiyoti universiteti



Samarqand davlat  
universiteti



Toshkent axborot  
texnologiyalari universiteti  
Samarqand filiali

## MUHAMMAD AL-XORAZMIY NOMIDAGI TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI SAMARQAND FILIALI

### “O‘ZBEK TILINING MILLIY KORPUSI: MUAMMOLAR VA VAZIFALAR”

MAVZUSIDAGI XALQARO  
ILMIY-AMALIY ANJUMANI MA‘RUZALAR TO‘PLAMI  
2023 yil 25 mart

PROCEEDINGS  
of the International Scientific and Practical Conference  
“THE NATIONAL CORPUS OF THE UZBEK LANGUAGE:  
PROBLEMS AND TASKS”  
March 25, 2023



Mono	295	8	93.798	95.451
Mono	295	16	95.412	96.97

Suxandonni tanib olish tizimini ishlab chiqishda ovoz namunalarning sonini ko‘p bo‘lishi hamda GMM algoritmining gauslar soni 16 ta bo‘lganda tizim anilik ko‘rsatgichi 96,97 foizga yetdi. Keying bosqichlarda tanib olishning gibrid algoritmlaridan foydalanish ishlari amalga oshirish rejalashtirilgan.

#### Foydalanilgan adabiyotlar ro‘yxati:

1. Mahboob, Tahira & Khanam, Memoona & Khiyal, Malik & Bibi, Ruqia. (2015). Speaker Identification Using GMM with MFCC. International Journal of Computer Science Issues. 12. 126-135.

2. M S, Sinith & Salim, Anoop & Sankar, K. & Narayanan, K. & Soman, Vishnu. (2010). A novel method for Text-Independent speaker identification using MFCC and GMM. 292 - 296. 10.1109/ICALIP.2010.5684389.

3. Mamyrbayev, O., Mekebayev, N., Turdalyuly, M., Oshanova, N., Ihsan Medeni, T., & Yessentay, A. (2020). Voice Identification Using Classification Algorithms. Intelligent System and Computing. doi:10.5772/intechopen.88239

4. K. Shukurov, T. Boburkhon and U. Khasanov, "Implementation of speech processing algorithms based on Singular Value Decomposition and Hidden Markov Model," **2021 International Conference on Information Science and Communications Technologies (ICISCT)**, 2021, pp. 01-03, doi: 10.1109/ICISCT52966.2021.9670357.

5. U.K. Xasanov. Shaxsni nutq orqali tanib olish usullari// “Intellektual axborot va kommunikatsiya texnologiyalari sohasidagi so‘nggi yutuqlar” mavzusidagi xalqaro ilmiy-amaliy anjumani “ISPC-2022”, Toshkent,23-24 Noyabr,79-82b

## THE ROLE OF DIGITAL AND INFORMATION TECHNOLOGIES IN AUDIOVISUAL DICTIONARY CREATION AUDIOVIZUAL LUG‘AT YARATISHDA RAQAMLI VA AXBOROT TEXNOLOGIYALARINING O‘RNI

*Shirinova Mekhrigiyo Shokirovna*

Bukhara State University, Bukhara, Uzbekistan

[m.sh.shirinova@buxdu.uz](mailto:m.sh.shirinova@buxdu.uz)

**Abstract.** In this article, the importance of digital and information technologies in the creation of audio-visual dictionaries related to linguistics is emphasized, and the ideas are proved on the example of "Orphoepic audiovisual dictionary of the Uzbek language (letters X and h)". It has also been proven that the use of digital and information technologies in the creation of audiovisual dictionaries creates all-round comfort for dictionary users and creates economic savings.

**Annotatsiya.** Ushbu maqolada tilshunoslikka oid audiovizual lug‘atlarni tuzishda raqamli va axborot texnologiyalarining ahamiyati yuqori ekanligi ta’kidlanib, “O‘zbek tilining orfoepik audiovizual lug‘ati (X va H harfi)” misolida fikrlar dalillangan. Shuningdek, audiovizual lug‘atlarni yaratishda raqamli va axborot texnologiyalardan foydalanish lug‘atdan foydalanuvchilar uchun har tomonlama qulaylik yaratishi, iqtisodiy tejamkorlikni yuzaga keltirishi kabi jihatlari isbotlangan.

**Keywords:** audiovisual dictionary, digital and information technologies, orthoepy, orthoepic dictionary, orthoepic standard.

**Kalit so‘zlar:** audiovizual lug‘at, raqamli va axborot texnologiyalari, orfoepiya, orfoepik lug‘at, orfoepik me‘yor.

Today, digital and information technologies have developed so much that, as a result, there are almost no areas where they have not penetrated. In particular, the field of linguistics is not an exception. Traditional methods of teaching with the help of textbooks and teacher’s explanations are no longer enough to develop students’ ability to think creatively, critically, search for scientific information, and analyze. It is necessary to use technologies that allow the implementation of modern approaches to education. When using digital technologies in the classroom, the motivation to learn increases, the cognitive interests of students are stimulated, and the efficiency of independent work increases. The computer opens fundamentally new opportunities in the field of education, in the student’s educational activity and creativity. In order to develop students’ cognitive activity, it is necessary to use educational technologies correctly, to use them not only in class, but also in extracurricular activities.

Digital technology is faster than any innovation in our history is developing. In just two decades, digital technologies have reached almost half of the population of the developing world and changed the way society lives. Digital technologies have emerged as systemic equalizers by enhancing access to finance, commerce and public services. For example, in healthcare, AI-based technologies are helping to save lives, diagnose diseases, and extend life expectancy.

In education, electronic and distance learning is helping to create curricula for students with physical disabilities. Digitization is making public services more accessible and transparent, which is evidence that digital technologies are helping to reduce barriers and increase efficiency in various sectors. Problems in the process of working with large amounts of data are also being solved as a result of technological development. All of these special recognition as progressive indicators of the development of digital technologies will be done.

There is a lot of work in Uzbekistan regarding the development of digital technologies being implemented. In particular, if we look at the following thoughts of the President: "We need to develop a national concept of Digital Economy, which involves updating all areas of the economy based on digital technologies. On this basis, we need to implement the program "Digital Uzbekistan-2030". Also, Sh. Mirziyoyev emphasized that special attention should be paid to information security when developing the concept of "Transition to Digital Economy". The adoption of the decision PQ-3832 by the head of our state "On measures to further modernize the digital infrastructure in order to develop the digital economy" will create conditions for the rapid development of the digital economy, further improve the state management system, expand the possibilities of its use and It is an important program in the application of modern infrastructure [1].

The issue of digital transformation of education remains one of the most relevant issues recently. For example, factors such as changing the principle of interaction with open information, increasing its availability or increasing its quantity lead to the globalization of scientific and technical culture, and also require the need to modernize a modern educational institution. In particular, the modern conditions of the surrounding reality require constant improvement of educational technologies by mutual adaptation of the pedagogical process and its tools. Undoubtedly, excessive conservatism, ignoring the widest possibilities of modern digital and multimedia technologies is equivalent to bringing the education system to a dead end. Therefore, the digital transformation of education in one form or another, whether it is a transition to a digital school or a transition to an information technology system, is a goal-directed process of qualitative change in the educational system through the mutual adaptation of didactics. and in order to improve the effectiveness of this process, teaching methods are inevitable.

Digital technologies are technologies that provide information in a "digitized" or universal digital form. Digital technologies are electronic tools, systems, devices and resources that enable the creation, processing, storage and transmission of information. Digital technologies in education

is the creation of an educational environment based on digital technologies. Today, the following digital technologies can be used in the educational process:

**Mobile technology** is a set of methods and tools used by users to exchange information through mobile devices.

**Cloud technologies** - the process of using computers on the Internet to store, manage and process data. Cloud technologies usually provide the user with computer resources, programs and power in the form of an Internet service. Cloud technology resources are located on several servers and are managed centrally.

**Smart technologies** are technologies that are created by humans and add additional functions to the traditional technology that performs a specific task, enabling the implementation of a number of mental activities. Including smart watch, smart suit, smart umbrella and others. In general, smart technology relies on smart devices equipped with one or more sensors that analyze a specific task based on pre-programmed rules, draw conclusions, and perform a task by communicating with other smart devices.

**The Internet of Things (IoT)** refers to the billions of physical objects or "things" connected to the Internet, all of which can collect and exchange data with other devices and systems over the Internet. Currently, many devices in a typical home environment have modern hardware (microcontrollers) and software. IoT seeks to make these devices work together and help each other to create a network of intelligent things. For example, if a microwave oven, a refrigerator, a door lock, a video surveillance camera or other similar devices are connected to the Internet, they can be controlled remotely using a smartphone.

**Virtual reality (Virtual reality - VR)** - an artificial information environment consisting of a set of software and technical tools that allow information to be conveyed to the user's mind using the senses (sight, hearing and touch). Virtual reality is an interactive technology that allows you to create the illusion that a person is moving in real life on a computer. In this case, the perception of the objective existence with the help of natural senses is replaced by artificially created computer information with the help of a special interface, computer graphics and sound. Neurons in the human brain respond to virtual elements in the same way as real-world elements. Therefore, a person perceives the virtual environment and reacts to the events happening in the virtual world as if they were happening in reality.

**Augmented reality (Augmented reality - AR)** is an information environment in which interactive digital elements are placed on the surrounding real world objects. This technology aims to fill the real world using virtual elements through a visual device. AR virtual objects are simultaneously populated and tracked with information from the camera to view the real world. As a result, the real world is filled with artificial elements and new information. This applies to conventional smartphones and tablets, augmented reality glasses, stationary screens, projection devices and other technologies. Can be done using programs.

**E-learning** is a method of organizing education through digital technologies, which can be done online or offline through a computer or other digital device. It can be accessed from most electronic devices such as computers, laptops, tablets or smartphones, providing a versatile and easy way for students to learn wherever they are.

**Management System (LMS)** is a human-machine complex or remote computer designed to manage learning activities consisting of a set of educational materials such as video lessons, lecture materials, presentations, books, and can work in communication mode. a form of 'lim.

Massive **Open Online Course (MOOC)** is a platform that allows students to study for a year or in short sessions. Through MOOC platforms, the world's most prestigious educational institutions, such as Harvard and Massachusetts, offer online courses in various fields and at various levels. It also allows thousands of students to register at once, to study as a member of the course during a fixed period of time or at any time throughout the year.

**Content management systems** (Content Management System - CMS) is a platform consisting of ready-made template structures, a set of functions for data entry and design. CMS is a site management system that allows you to easily and easily create and modify (add, edit and delete) sites consisting of text and multimedia pages.

**Big data** (Big Data) is a collection of data that grows over time and includes technical and software tools aimed at performing actions on the data. Big data offers the ability to collect, process and transmit large amounts of data, unlike traditional databases. It helps to obtain the necessary information through effective data analysis.

**Knowledge base** (Data Science) is an interdisciplinary field, which is usually aimed at obtaining new knowledge, that is, useful information, from a large amount of data. This area determines the preparation of data for analysis and critical information needed for high-level decision-making in the organization.

**Artificial intelligence** is a system with the ability to correctly interpret external information, learn it and use this information to achieve specific goals and tasks by adapting to different situations. Artificial intelligence allows machines to imitate human intellectual behavior and thinking.

**An infographic** is a graphical way of presenting information, simply put, information presented in visual form. Infographics use visual means of information representation such as graphics, images, diagrams, tables, maps, schemes. Infographics are designed to present information quickly, visually and vividly.

Orthopedics is a branch of linguistics that studies the rules of correct pronunciation. Correct pronunciation is the crown of every language. Just as a person without clothes is a laughing stock, a word that is not properly pronounced is a laughing stock not only of a language but also of a nation. Pronunciation techniques play a key role in learning any language. Let's say that when the representative of each nation starts speaking in his native language, he is not yet literate and has not mastered theoretical information. He learns words only based on the pronunciation of people around him. This appropriation occurs in different forms and dialects. What is right or wrong is defined by the orthoepic norm. Why is this limit necessary? First, uniformity in pronunciation is ensured. Secondly, it calls the nation to unity. Thirdly, the language learner avoids various confusions that occur among representatives of other nations.

Pronunciation standards should be taught step by step. Because a pronunciation dialect adapted to a certain dialect cannot be taught to pronounce correctly at once. The best helper for us in this regard is a dictionary. There are several tasks that need to be carried out in the field of lexicography of the Uzbek language. This is confirmed by the following evidence: the 1st pronunciation dictionary in Uzbek was compiled in 1977 by M. Sodikova and O'Usmonova. 7 years later, in 1984, a number of co-authors under the leadership of M. Asomiddinova created the "Uzbek Literary Pronunciation Dictionary". After that, not a single dictionary was created within the framework of orthoepy. However, if we look at the Russian lexicography, we see a variety of dictionaries related to the same pronunciation, and these dictionaries have not been published for a long time. We also see grouping by age (eg separate and different pronunciation dictionaries for preschoolers and school-age).

In our opinion, it is better to choose an audio or audiovisual dictionary form, since orthoepic dictionaries teach correct pronunciation according to their content. Because during hearing, exactly how to pronounce the word is clearly heard, understood and mastered. According to the statistics of Ruiperez, a person learns through 10% audio (i.e. hearing) sources, 20% visual (i.e. seeing) sources, and finally 65% audiovisual (i.e. hearing) sources. studies [2]. So, in acquiring knowledge, seeing and hearing the source at the same time gives an effective result.

If you refer to the dictionaries named above, the correct pronunciation of the word is expressed in the transcription tool. Understanding and understanding it belongs only to experts.

Sometimes even philologists may not understand. So, how should an ordinary student or a reader who wants to learn the language get information?!

The dictionary we would like to recommend to you is "Orphoepic audiovisual dictionary of the Uzbek language (letter X and H)". This dictionary contains only the active words in Uzbek that contain the letters x and h. The dictionary is designed for the site. The words are given alphabetically in Latin and Cyrillic scripts. If you want to know the pronunciation of a word, place the cursor over that word and press the char button of the mouse, or you can quickly find it by typing the desired word in the search field. When you press the cursor, you will hear the pronunciation of the word. If you want to hear it again, you press the cursor again. This dictionary also replaces a spelling dictionary because you can see how the word is spelled.

This dictionary can be used for the following purposes:

1. **In the formation of correct pronunciation in preschool children.** In this case, the educator can use it in the process of working on the speech of children who do not distinguish between x and h, or when teaching the pronunciation of any word related to the topic.

2. **In the formation of correct pronunciation in school-aged children.** It is known that school is the most important stage of education, because 9 or 11 years of a person's life are spent in school. During this period, a person fully develops his pronunciation standards. It would be an excellent light upon light if the teacher taught the pronunciation of each new word he was teaching by referring to the dictionary.

3. **In forming the correct pronunciation of representatives of other nationalities who want to learn the Uzbek language.** Today, interest in the Uzbek language is increasing day by day. Therefore, it is necessary to provide comfort and ease in language learning. We believe that this dictionary will help in this regard. It also prevents various confusions in the pronunciation of words, making it convenient for independent language learners.

**In forming the correct pronunciation of the general Uzbek-speaking public.** It is important to teach correct pronunciation to the general public, i.e. representatives of other fields who are not philologists, managers and workers.

In conclusion, it can be said that it is necessary to improve Uzbek pronunciation dictionaries. After all, "Orphoepic audiovisual dictionary of the Uzbek language (letter X and H)" contains only more than 3000 words. However, the Uzbek language has more than 100,000 words and phrases. One of the main future tasks is to create a comprehensive "orphoepic audio-visual dictionary of the Uzbek language". True, the compilation of this dictionary requires long and hard work, as well as funds. Also, this work cannot be done by just one person. If we want to spread the Uzbek language widely, promote it, and make it to the world, it is necessary to compile this dictionary together with many people, because we compiled the "Orphoepic audiovisual dictionary of the Uzbek language (X and the letter H)" is patir from the end of the dough. It should also be noted that this dictionary serves as material for several audiovisual dictionaries that will be created in the future.

#### **List of used literature:**

1. O‘zbekiston Respublikasi Prezidenti Shavkat Mirziyoyevning 2020- yil 24- yanvarda Oliy Majlisga yo‘llagan Murojaatnomasi. <https://uza.uz/oz/politics/zbekiston-respublikasi-prezidenti-shavkat-mirziyeevning-oliy-2501-2020>.

2. Юнусова Д.М. Роль электронного словаря в современном мире//Лексикография ва фразеологиянинг долзарб муаммолар. – Т., 2014. –Б. 74

3. R.X.Ergashev, S.N.Xamrayeva. Raqamli iqtisodiyot - (Darslik). – Т.: “Intellekt”, 2021. – 475 b.

4. Fayziyeva M.R. Ta’lim transformatsiyasida raqamli texnologiyalarning ahamiyati. Science and innovation. International Scientific Journal Volume 1 Issue 8 UIF-2022: 8.2 | ISSN: 2181-3337

### 3-SHO‘BA. TILSHUNOSLIKDA RAQAMLI VA AXBOROT TEXNOLOGIYALAR

63. **Мусаев М.М., Очилев М.М., Хўжасяров И.Ш.**  
Нутқни таниш тизимни яратиш учун нутқ корпусини ҳосил қилиш муаммолар ва ечимлари 247
64. **Elov B.B., Axmedova X.I.**  
Turli so‘z turkumlari orasidagi omonimiyani na‘ive bayes kalsifiktori yordamida aniqlash 251
65. **Вафоев М.А., Толибов С.В., Норбеков Р.И.**  
Лингвистическая стегосистема для сокрытия коротких сообщений в узбекском тексте 257
66. **Tursunov J.**  
Incorporation of natural language processing in IOT projects 261
67. **Xolmatov O.A., Kamolov R.K.**  
NLPda savol-javob tizimlarini yaratish turlari va bosqichlari 266
68. **Bobamurodova S.E.**  
The usage of ict and digital technology in language learning and linguistics 270
69. **Djumayev S.N., Ortiqova N.E., Asrorova M.G‘.**  
Lotin alifbosidagi imloviy xatolarni tekshirishga mo‘ljallangan mobil ilova tavsifi 273
70. **Naziraliyev U.**  
Natural language processing techniques for speech signal analysis 276
71. **Kaxxarov A.A., Shukurov K.E., Xoldorov Sh.I.**  
Nutqni tanib olishni amalga oshirish bosqichlari 281
72. **Iskandarova S.N., Maxkamova D.A.**  
Forscha matnlarni o‘zbekchaga tarjima qilishda NLP(natural language processing) dan foydalanish samaradorligi 285
73. **Uzoqova M.T.**  
O‘zbek tili sinonimayzeri xususida 288
74. **Mahmudjonova G.U.**  
TF-IDF asosida o‘zbek tili matnlarini tasniflash 292
75. **Samatboyeva M.T.**  
NER: o‘zbek tilidagi matnlarda nomlangan obyektlar tavsifi va ularni avtomatik aniqlash 297
76. **Shukurov K.E., Xasanov U.K., To‘rayev X.Sh.**  
Nutq signallarini orqali suxandonni tanib olish 301
77. **Shirinova M.Sh.**  
The role of digital and information technologies in audiovisual dictionary creation 305
78. **Suyunova M.O.**  
O‘zbek tilida matnni avtomatik tahlil qilish 310
79. **Negmatova N.E.**  
Tilshunoslikda raqamli va axborot texnologiyalari hamda kompyuter lingvistikasining ahamiyati 314
80. **Mamatova D.M.**  
Tilshunosligimizda raqamli dasturlash muammolari 319
81. **Shodmonov D.A. Djumayev S.N.**  
Ish yuritish hujjatlarini tayyorlashni avtomatlashtiruvchi axborot tizimini yaratish bosqichlari 323