UDC 37:004.92

Olesia Barkovska
PhD of Computer Science, Associate professor, Department of Electronic Computers
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID 0000-0001-7496-4353
olesia.barkovska@nure.ua

Heorhii Ivashchenko
PhD of Computer Science, Associate professor, Department of Electronic Computers
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID 0000-0003-1027-5262
heorhii.ivashchenko@nure.ua

Dmytro Rosinskiy
Senior Lecturer, Department of Electronic Computers
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID 0000-0002-0725-392X
dmytro.rosinskyi@nure.ua

Daniil Zakharov
Master's student of Department of Electronic Computers
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
ORCID ID 0009-0006-4676-6599
daniil.zakharov@nure.ua

EDUCATIONAL TRAINING SIMULATOR FOR MONITORING READING TECHNIQUE AND SPEED BASED ON SPEECH-TO-TEXT (STT) METHODS

Abstract. The work is devoted to a relevant issue – gamification of the educational process of primary school students with the use of digital mobile devices using the example of educational trainer developed to control the technique and speed of reading. The practical novelty of the study lies in its potential to optimize the process of controlling reading techniques for children with various speech disorders, such as lisping, rhotacismus, and dyslalia, which remains an unsolved problem for existing computer linguistic models. The practical significance lies in the fact that the use of information and communication technologies and gamification of the learning process motivates to gain knowledge and encourages learning not only due to the game strategy, but also due to the fact that the child receives approval and the opportunity to perform the task again so to improve skills. The proposed and tested trainer for monitoring the technique and speed of reading of primary school students for a mobile digital device recognizes the child's voice with high accuracy (up to 94.84%), detects misread words and counts the number of words read in the given time period. The architecture of the proposed system consists of the following modules: preparation module, which ensures the correct functioning of the system regardless of the existing speech disorders; the speech-to-text converter directly converts the voice of the child reading the text on the gadget screen into text; The text comparison module is responsible for comparing the text read aloud with the text obtained as a result of conversion. The results of testing children with lisping and rhotacismus problems showed that the Jaro algorithm has a slightly higher accuracy of comparing two texts (by an average of 1.14%), and, at the same time, a shorter comparison time for large text arrays (by 39%). The results of testing children with dyslalia, meaning the rearrangement of sounds in words when reading, also showed that the Jaro algorithm has a shorter line comparison time for large text arrays (by 7%). The reduced operation time in methods for comparing texts in dyslalia cases is attributed to the absence of text array variations, with evaluations solely focusing on potential permutations of sounds.

Keywords: computer game; reading technique; mobile devices; editorial distance; speech-to-text.

1. INTRODUCTION

The problem statement. Modern conditions of the education system require teachers to possess qualities such as mobility, creativity, practical orientation, and others. However, perhaps in the age of digitalization, one of the necessary conditions for successful pedagogical activity is a high level of information culture. One of the trends in education is to encourage the allowance for mistakes. Therefore, it is crucial in the application not to highlight task failures, whether through character dialogues or sounds, and instead provide opportunities for hints or explanations. It is also important for development to be able to replay the level to correct mistakes.

Thus, the use of a computer program increases motivation not only due to the game strategy on which the program is based, but also because the child receives approval and praise not only from adults, but also from the computer [1].

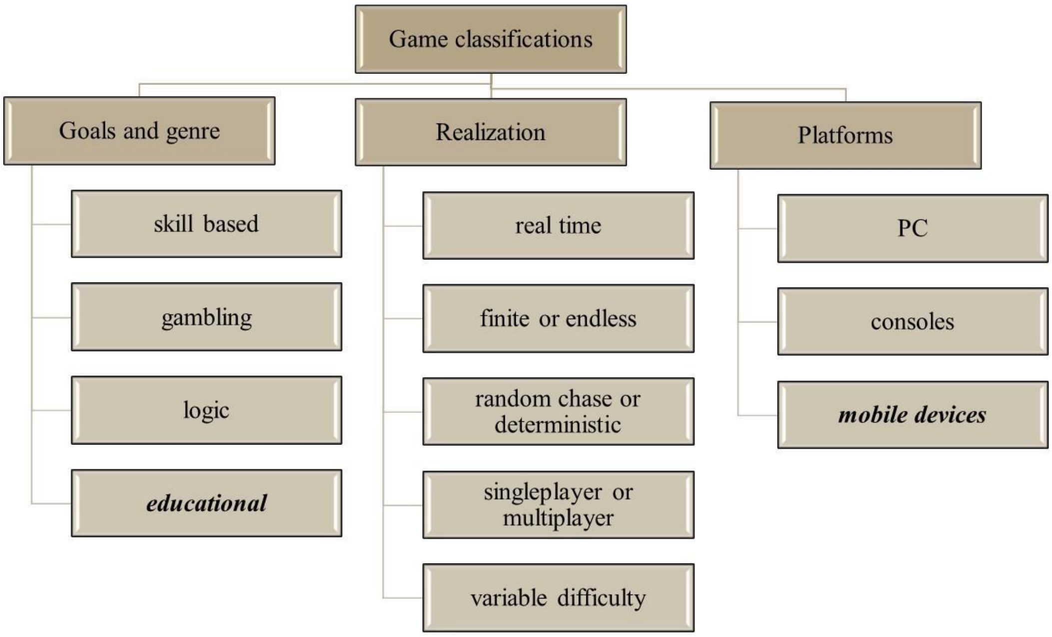

Figure 1. The place of educational games on mobile devices in the general classification of games

A modern smartphone or tablet can completely replace all the functionality of home computers. In this regard, the development of educational applications is becoming relevant not so much for personal computers but for mobile devices (Figure 1) [2], [3]. It should be noted that both specially designed apps and web-based implementations can operate on mobile devices.

Analysis of recent studies and publications. Digital learning is an educational practice that helps students and yields tangible results [4] – [8]. The use of information and communication technologies has become an obligatory part of the educational process. Making learning interesting is facilitated by such a development as gamification – the use of gaming rules and approaches to achieve educational goals [9]. In other words, boring tasks become interesting, presented material becomes desirable, and complex becomes simple. Trends in the development and prospects for the growth of the mobile gaming market are shown in Figure 2.

However, the complete reliance on interactive educational materials is not acceptable, as/since gaming platforms have disadvantages highlighted by psychologists:

– lack of communication and social aspects of learning, such as information exchange and emotional involvement in the process;

– insufficient level of control in terms of discipline.

To prevent negative consequences, it is necessary to combine gamification techniques with classical methods. Gamification should not play the main role but rather serve to motivate students to gain knowledge [10], [11].

The statistics shown in Figure 2 confirm that the development of mobile games is now in its most active phase, which confirms the relevance of research and improvement of methods used in a number of educational applications for mobile devices. The mobile gaming market size is expected to grow from $141.71 billion in 2023 to $300.47 billion.USD by 2028 at a CAGR of 16.22% during the forecast period (2023-2028).

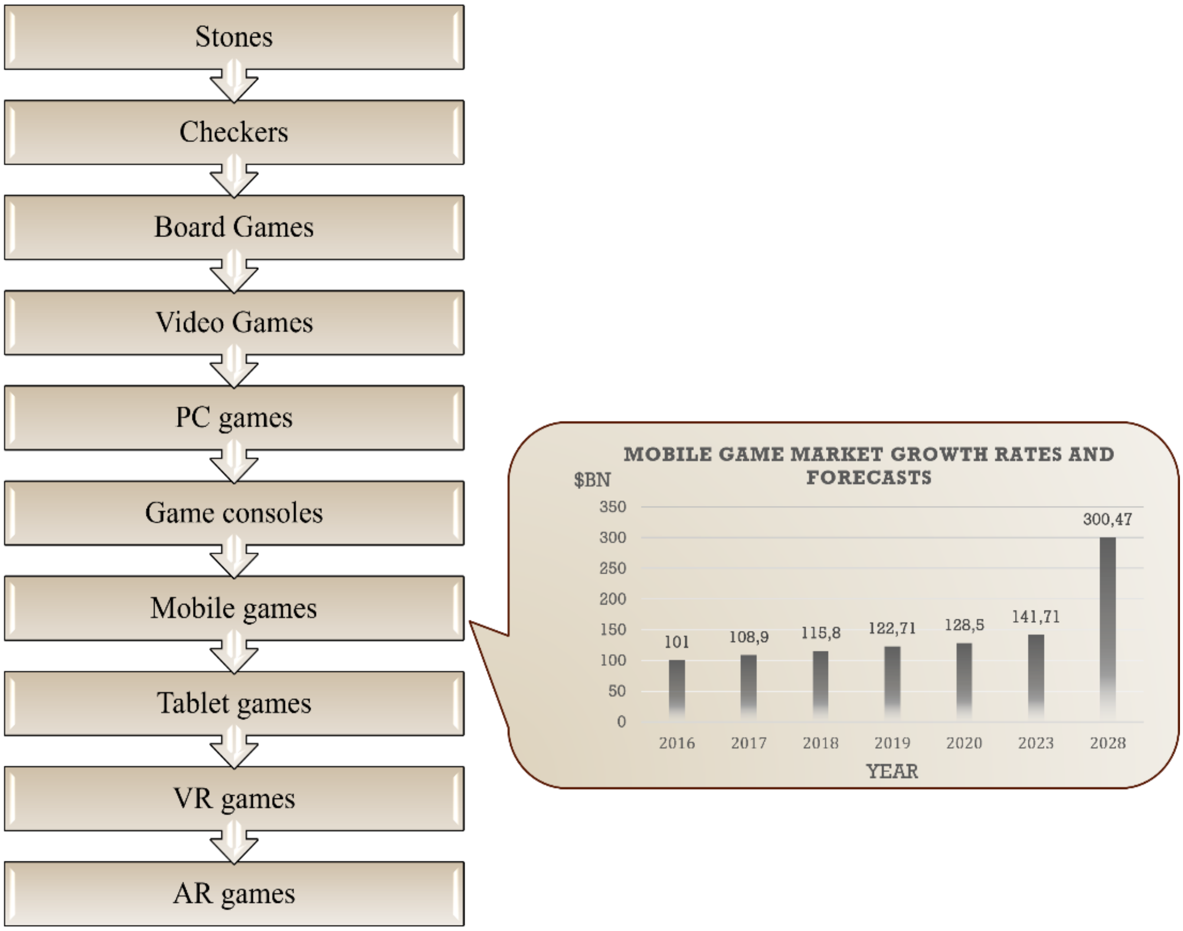

Figure 2. Trends and Growth Prospects of the Mobile Gaming Market

The diagram is provided to show the place and demand for educational games on mobile devices, as a justification for the relevance of the work performed.

When developing interfaces for children's gaming applications, one may encounter the following general recommendations: a straightforward interface, vivid graphics, sound accompaniment for every action within the game, the ability to replay levels [12] (see Figure 3).

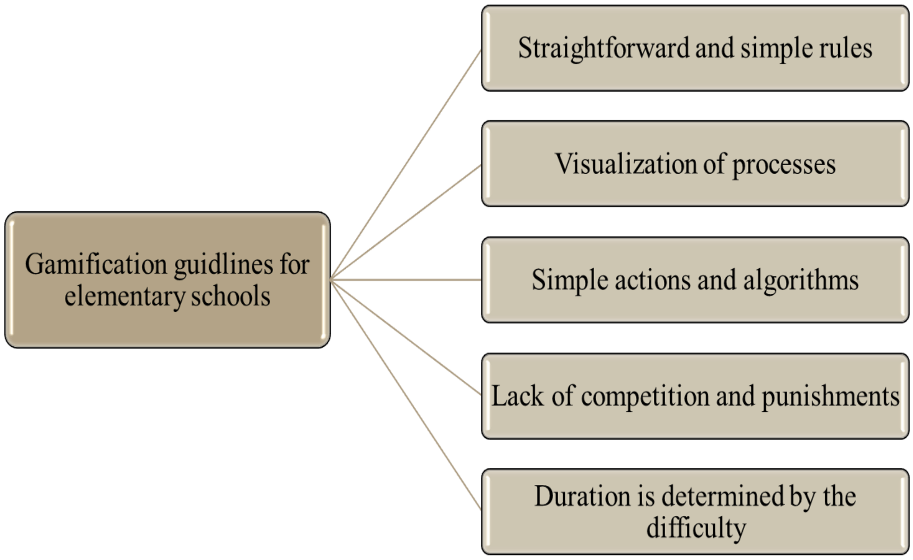

Figure 3. Gamification rules for primary school

The most important advice is to focus on the age and characteristics of the child's development during this period [13], [14]. First of all, you need to ask yourself the following questions: whether the user can read, whether he can play the game on his own, whether parents should perform tasks with him. Developers of children's games believe that users play the app the same way they interact with the world around them. In the outside world, there is a reaction to any of our actions, for example, a sound or the movement of other objects. And in the game, the child will wait for the program's reaction to his actions. Therefore, it is better to accompany the entire game with sounds, ensuring that even button presses produce a notifying sound.

The research goal. The objective of this work is to develop an educational training simulator for monitoring the reading technique and speed of elementary school students on a mobile digital device. To achieve this goal, the following tasks must be addressed:

- analysis of requirements and principles for creating a mobile game for children, taking into account the age-specific characteristics of the target audience;
- development of a model for an automated reading technique monitoring system for young school-aged children;
- investigation of Speech-to-Text (STT) services based on criteria such as accessibility, dependency on internet connectivity, and computational resource requirements;
- study of methods for determining string similarity based on character-level comparison;
- analysis of the obtained results.

The research object is an educational application for mobile devices that can be used during the educational process of elementary school students to monitor reading technique and speed. The research subject is the STT methods applied to analyze the text read by the child.

The proposed educational game should recognize the child's voice, accurately detect incorrectly read words, and accurately count the number of words read within the allotted time frame.

2. RESEARCH TASK RATIONALE

Reading is a key position of successful learning for any student. Experiments conducted in recent years have shown that speed reading activates thinking processes and is one of the means of improving the educational process for different levels of education, from primary to high school.

Optimal reading is at the speed of spoken language, that is, at a pace of 120 to 150 words per minute. That is the speed that the human articulatory apparatus has adapted to over many centuries, it is at this speed that a better understanding of the text is achieved.

Every day, new manuals, games, and recommendations emerge, which eventually lead to the achievement of the goal – a child begins to read. Next, it is necessary to teach the child to read quickly, expressively, and with a good comprehension of the text, thus developing "reading technique".

During the reading technique test, the following are evaluated:

- reading speed – the number of characters a child can read per minute;
- reading method – whether the child reads words syllabically or as whole units, smoothly;
- accuracy – the absence of errors and hesitations made by the child while reading;
- comprehension – the ability to understand the meaning and idea of what has been read;
- expressiveness – the skill to correctly place emphasis, follow intonation, and maintain pauses during reading.

Analyzing the listed criteria, we can say that the reading technique test is based on the assessment of two components: content and technique. Meaning, the technical side – tempo, expressiveness, correctness – is subordinated to the content, that is, the ability to understand the meaning of the text.

Thus, the application under development is a training simulator to improve technique and reading speed.

AI methods and technologies for speech-to-text (STT) and text-to-speech (TTS) conversion are essential for creating effective, accessible, and inclusive educational simulators not only for monitoring and training reading skills but also for gamifying the learning process in general, because they:

- promote personalized learning, increasing interaction and engagement among children;
- provide support for children with special educational needs;
- enhance the efficiency of the learning process and contribute to the development of technical and communication skills.

These technologies make them indispensable tools in modern education.

Integration of voice commands and feedback makes educational simulators more game-like (gamification of the learning process), helping children better absorb material in an environment they perceive as entertainment.

The use of AI methods and STT and TTS technologies [15], [16] allows for the adaptation of content and teaching methods, ensuring the optimal development of technical skills and reading speed. This includes adapting educational content to the linguistic characteristics of the child, which is crucial for children with different accents, dialects, or speech impairments.

Computational and artificial intelligence methods, which utilize artificial linguistic models, are actively used in modern educational applications and simulators. They are adapted for the needs of people of different ages as well as for the tasks set before the application. These range from learning foreign languages [17], virtual training avatars for gaining experience in interviewing children subjected to abuse [18], controlling the impact of AI methods on critical thinking skills [19], language competence skill enhancement simulators for learning foreign languages [20], in systems of testing and monitoring students' knowledge, which are part of LMS [21] and others.

Among the reviewed applications available on the Ukrainian market and designed to improve the speed and technique of reading, the following were considered: Readlax, Speed

Reading – Brain Development, Reedy – Intelligent reader, Reading Trainer, Balto Speed Reading (Table 1).

Readlax is an app for people who want to read faster and with the better comprehension. Readlax is an online brain workout for speed reading. The main goal is to help you gain knowledge faster, improve your reading speed and improve your reading speed and comprehension.

Readlax has developed two products: speed reading games and a chrome extension with integration of Google Books, Scribd, and Kindle Cloud Reader.

The developers of the Speed Reading – Brain Development app have combined training exercises with the ability to read books directly in the app. All workouts are divided into blocks and levels. The advantage of the application is its training course. It will include blocks for memorization, expansion of the visual window, training of the point of concentration and reading speed.

In Reading Trainer, there are built-in games that are aimed at training quick recognition of numbers, letters, and words; the user can train free eye movement; increase the ability to concentrate and expand visual range. The developers have built personal user statistics into the app, where you can evaluate your own progress.

The English-language Balto Speed Reading application implements the functionality of a trainer and a reader of all formats. At its core, this is a reader with advanced functionality for training speed reading.

Table 1

Comparative analysis of existing analogues

Application	Advantages	Disadvantages
Readlax	- 7-day trial period - exercises are designed in the form of short games with gamification integration (points, achievements, badges) - the Chrome extension has integration with Google Books, Scribd, Kindle Cloud Reader - the ability to choose difficulty levels - training in different languages - it is possible to upload your own text for training - speed testing and progress monitoring - simplicity and accessibility of the interface	- subscription cost after the trial period is $9.99 USD per month - there is no player rating
Speed Reading – Brain Development	- free access - the ability to train using books from your own library - training available in books of any language - variety of exercises	- lack of player rating - it is not possible to choose your level
Reading Trainer	- the ability to track your own progress - available in 10 languages	- paid access - unintuitive and overloaded interface - lack of Ukrainian language
Balto Speed Reading	- manual navigation of the e-reader	- lack of Ukrainian language - primitive interface

	- the ability to load a large number of formats and web pages - intelligent division into paragraphs for ease of reading	- all reading control options are available with the purchase of the paid version

After reviewing the available applications, we can conclude that none of them analyze reading technique, reading attentiveness and speed in the early stages, when the child is just learning to read, instead, they are aimed at improving acquired skills. This once again confirms the relevance of the topic of the research and the study of voice analysis methods.

To test reading technique, texts used should meet several requirements:

– they should be unfamiliar to the child yet age-appropriate and not too complicated;
– sentences should be short and free from dialogues or excessive punctuation;
– texts without pictures are preferable;
– font size should be large enough for easy readability;
– the text should not be divided across two pages.

Of course, the reading technique should not be tested using the instructions for household appliances or technical manual. The best option is short stories about nature, animals, and seasons. According to the requirements of the elementary school student academic achievement monitoring and assessment system, by the end of the first semester, students are expected to read the following number of words per minute: 2nd grade – 35-45 words per minute, 3rd grade – 65-70 words, 4th grade – 80-85 words.

During the development of the proposed reading technology control system architecture, the selection of the platform for further development and the software services available for said platform were justified. Comparing the characteristics of mobile operating systems Android, iOS (Apple), Windows Mobile (Microsoft), HarmonyOS (Huawei), and KaiOS is beneficial for determining user needs and is based on statistical indicators of market penetration internationally and in the European market [22], [23]. Android is the world’s largest operating system for mobile devices, supported by numerous data sources confirming its dominance. iOS holds the second position, although the difference in market share is notably significant. Android’s market share exceeds 70%, while iOS holds the second position with nearly 25% of the market share.

Based on the comparative analysis of Android, iOS, Windows Mobile, HarmonyOS and KaiOS mobile operating systems according to various criteria, the development will be adapted to the Android mobile operating system, as Android is an incredibly versatile platform that works on smartphones, tablets and a variety of embedded devices such as smartwatches and Internet of Things devices.

3. THE RESULTS AND DISCUSSION

During the development of architecture for the proposed reading technique control system, it was important to take into account not only speed, but also other criteria, such as the correctness of the text read. All errors made must be documented and analyzed during the process of correcting reading skills. Based on this, the following architecture is proposed, which consists of modules (Figure 4):

– preparation module, which ensures the correct functioning of the system, regardless of the existing speech disorders;
– speech-to-text converter directly converts the voice of the child reading the text on the screen into text;

– the text comparison module is responsible for comparing the text read aloud with the text obtained in the conversion result.

The described architecture is universal, suitable for different platforms, the difference lies in the internal organization and software used for a particular platform.

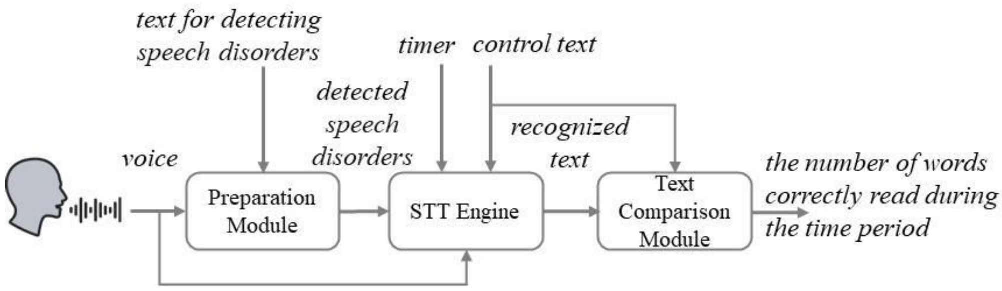

Figure 4. Architecture of the proposed system – simulator training simulator for control of technique and reading speed

When examining the functionality of the first module in more detail, it is worth noting which specific speech impairments are identified at this stage (see Table 2). The tasks selected at this stage aim to detect the accuracy of pronunciation of individual sounds. The necessity of identifying speech impairments before using the application arises from the fact that speech-to-text (STT) systems may have limited adaptability to various types of speech impairments. Enhancing the effectiveness of STT can be achieved through prior adaptation to specific types of speech impairments (such as stuttering, aphasia, dyslalia, background noise).

Immediately prior to assessing reading speed, the presence of pronunciation defects is examined using a series of examples presented sequentially on the screen.

Moving to the second stage – speech-to-text conversion, it is advisable to explore existing software services and their peculiarities, focusing specifically on their utilization on Android devices.

Table 2

Types of violations that are determined at the first stage (preparatory)

Type of violation	Manifestation of the violation	Is it fixed at the preparatory stage?
Articulatory disorders	problems with the clarity and accuracy of pronunciation of sounds (lisping, rhotacism)	yes
Dyslalia (dyslexia)	sound errors or incorrect arrangement of sounds in words; slow motion	yes
Tachylalia	speech that is too fast, which can affect clarity and comprehension	no
Stammering	interruptions in the rhythm and fluence of speech, such as repetitions, delays, or blocks	no

Let's highlight the following criteria for comparing STT services (Figure 5):

– access mode – online or offline;
– working principle – real-time speech-to-text or batch speech-to-text (streaming speech recognition and batch speech recognition).

Streaming Speech Recognition is used for real-time speech recognition, where text is transmitted to the server in parts, gradually, as the audio arrives. This approach allows you to get recognition results almost instantaneously, which is useful for applications such as voice assistants, recognition of control commands, and other scenarios where processing speed is important.

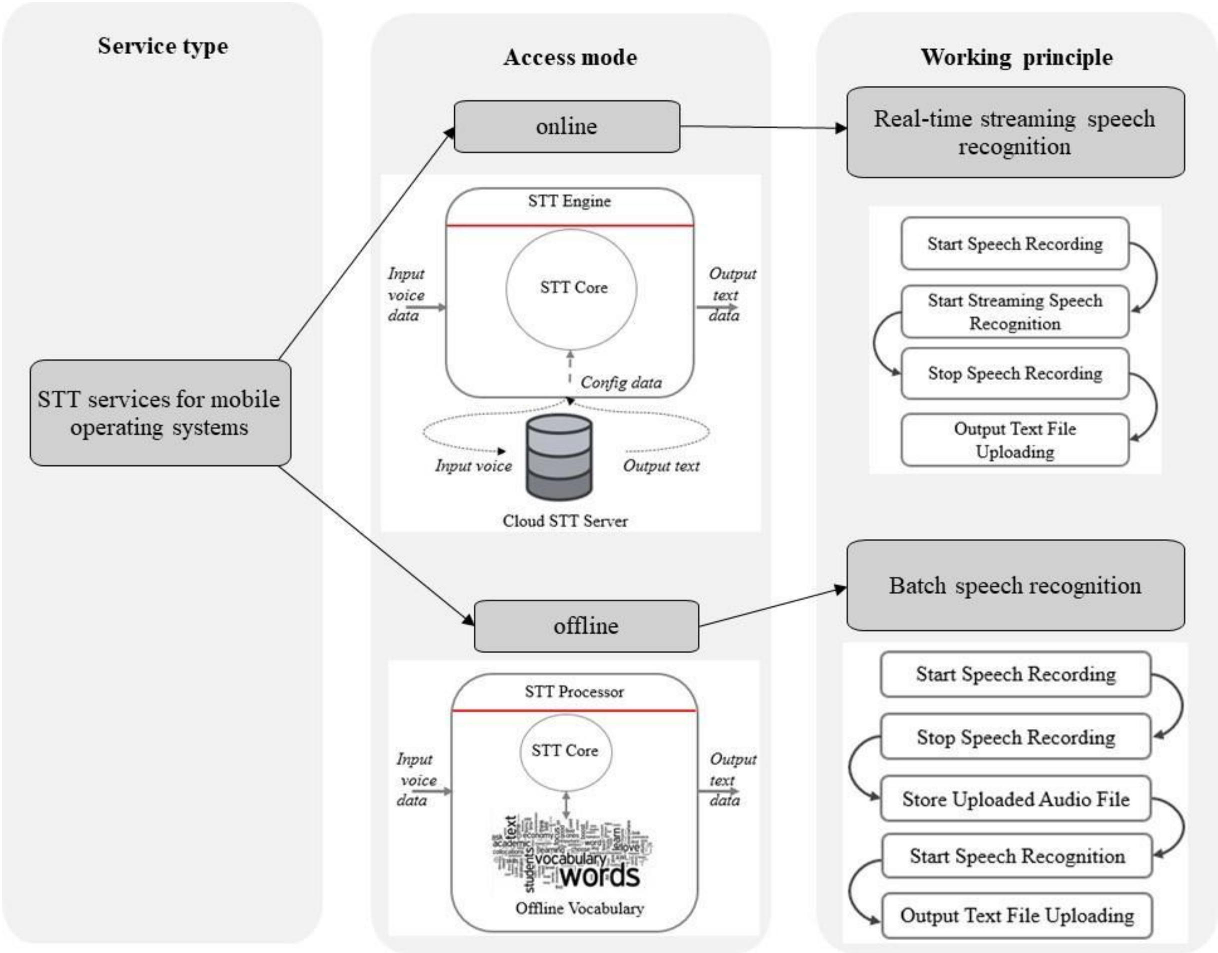

Figure 5. Overview of STT mobile platform services

Batch Speech Recognition is used to process audio files or large amounts of data in batch mode. Here, the audio files are completely transferred to the recognition system, and the results are obtained after the processing of the entire file is complete. This can be effective for large tasks or analyzing an audio archive, but it’s not as effective for real-time.

Based on the system’s requirements for the possibility of use in the absence of an Internet connection, an offline voice-to-text conversion approach was chosen, based on a batch voice-to-text conversion approach.

The last module is the Text Comparison Module, designed to perform character-by-character comparison of two texts while ensuring the precise consideration of word and character order. This task is closely related to substring search and textual similarity determination, crucial in information retrieval, document clustering, disambiguation, machine translation, and text summarization tasks, where measuring similarity between words, sentences, paragraphs, and documents is paramount.

There are many algorithms for determining string similarity – character-based and term-based. Examples of string similarity algorithms based on character-by-character comparison are:

- Longest Common Substring (LCS) – the algorithm of the longest common substring, takes into account the maximum length of a continuous chain of characters that exist in both lines;
- Levenshtein distance – determines the distance between two lines by calculating the minimum number of operations (insertion, deletion or replacement of one character or moving two adjacent characters) required to convert one line to another;
- the Damerau-Levenshtein distance is an extension of the Levenshtein distance, which further takes into account the transposition operation (permutation of two symbols);
- the Jaccard similarity index is used for sets of characters and defines the common part of characters in two texts. The similarity coefficient is calculated as the number of characters in common divided by the total number of unique characters in the two texts;
- Jaro similarity is calculated based on the number of characters shared and the number of transpositions between characters. This method is often used to compare texts that may contain errors or typos;
- the Needleman-Wunsch and Smith-Waterman algorithms are designed to compare biological sequences, but are successfully used to compare text strings by treating each string as a sequence of characters and creating a matrix in which the rows and columns correspond to the characters in your text to further populate the matrix, capturing matches and non-matches when comparing characters.

Table 3

Comparison Table of Levenstein and Jaraud Algorithms

Algorithm	Advantages	Disadvantages	Types of Operations for Calculating Editorial Distance
Levenshtein	- takes into account editing, insertion, and deletion of characters, allowing you to determine the exact number of operations required to convert one text to another - it can be useful if the number of edits is important to determine the similarity of texts	- doesn't always reflect similarity well, unless the number of edits is critical - requires computing resources, especially for longer texts	- remove - insert - replacement
Jaro Similarity	- takes into account the difference in character order and gives weight to the similarity of the sequence sought - takes into account the scalability of the length of texts, while less depends on the specific size of the text	- may be less sensitive to insertions and deletions, so it may not identify them as clearly	- permutation

By analyzing the existing algorithms of Levenshtein, Damerau-Levenshtein, Jaccard, and Jaro based on criteria such as the accuracy of comparing two texts, execution time, performance on large texts, performance on small texts, handling different languages, types of operations for calculating edit distance (including deletion, insertion, substitution, transposition), and the specificity of the existing task, the decision was made to investigate the methods of Levenshtein and Jaro (see Table 3).

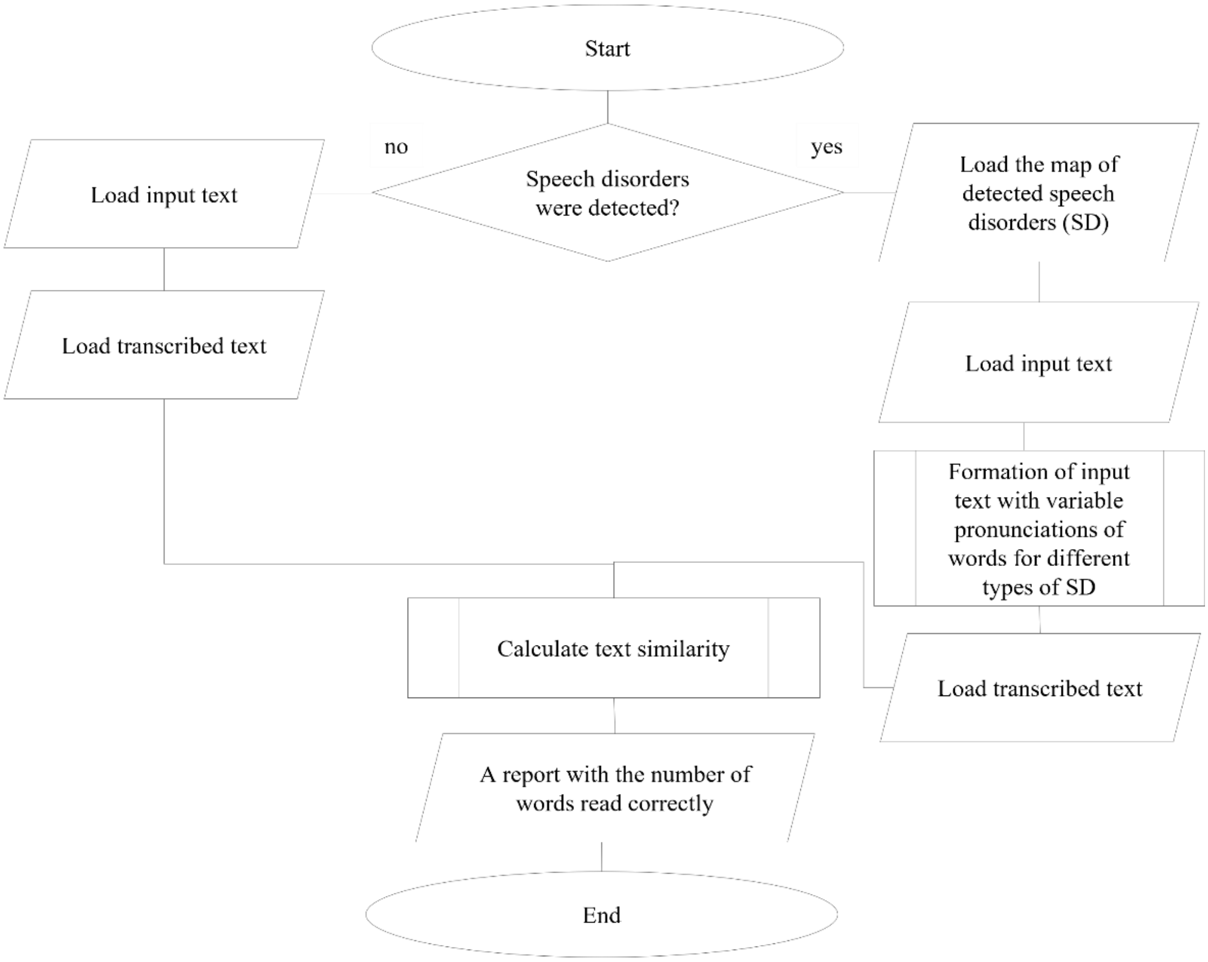

Figure 6. Algorithm of the system at the stage of comparison of text arrays

The general operational algorithm of the system, following the detection of speech impairments and the conversion of a child's voice reading the provided text, is structured as follows: upon identifying speech impairments, a map of these impairments is constructed, upon which variations of input text are formulated, taking into account various possible pronunciations of words. Subsequently, text comparisons are conducted. If no impairments are detected, the transcribed text is compared with the input. However, if impairments are identified, the transcribed text is compared with texts formulated considering the identified impairments (see Figure 6).

For the user, the system looks as follows: after adjusting the system to his voice and fixing the existing pronunciation defects or their absence, the user must enter the class in which the child studies to download the text of the appropriate complexity, click the "Start voice recording" button and start reading the text on the screen. After 1 minute, the child will be informed with an appropriate sound signal. After clicking on the "End Voice Recording" button, the page with the displayed text will be reloaded and the user will see a report with the number of correctly read words per minute and a list of words that were read incorrectly.

The evaluation of the methods used at the stage of converting voice to text was based on a comparison of the two approaches described above – online STT service (Google Cloud Speech API) and offline STT service (Android Speech to Text API (Speech Recognizer class). The results are shown in Table 4.

Analysis of the results showed that the use of the free service Android Speech to Text API (Speech Recognizer class) allows you to use the proposed system quite accurately with small time delays.

Table 4

Study of STT services according to the criteria of voice-to-text conversion speed, cost of use, requirements for computing resources, the need for Internet connection

		Online STT service Google Cloud Speech-to-Text	**Offline STT service Android Speech-to-Text**
Voice-to-text speed		almost instantaneously, depends on the quality of the Internet connection	delay up to 36 seconds
Voice Recognition Accuracy	no prior training	up to 87.6%	up to 81.9%
	with prior training	up to 96.7%	up to 92.3%
Cost of using the service		$0.009 per 15 seconds of audio	free
Compute Resource Requirements		has no requirements for the computing resources of a mobile device	from 5 MB per record of read text

The evaluation of methods employed during the stage of comparing text arrays was conducted across user categories with various speech impairments. The first category of children exhibited difficulties with lisping and rhotacism, while the second category of users experienced dyslexia-related issues, specifically involving sound permutation in words during reading. Consequently, the research results are presented in two independent tables.

The findings, outlined in Table 5, indicate that the Jaro algorithm demonstrates a slightly higher accuracy in comparing two texts (on average by 1.14%), while also exhibiting shorter comparison times for large text arrays (by 39%).

Table 5

Results of comparison of Levenstein and Jaro methods for a group of users who have problems with lisping and rhotacism

Text	**Algorithm**	**Average Comparison Accuracy**	**Average comparison time for all text variants generated**
TEXT A (60 words)	Levenshtein	93.47 %	2.25 s
	Jaro	94.84 %	2.37 s
TEXT B (89 words)	Levenshtein	91.3 %	2.49 s
	Jaro	92.05 %	2.9 s
TEXT C (135 words)	Levenshtein	90.47 %	3.36 s
	Jaro	91.56 %	3.04 s

The average accuracy of comparing two texts was calculated for all variants of texts generated based on the original text for specific speech disorders.

Sensitivity to permutations refers to the percentage of correctly identified permutations out of a total of 100%. Thus, if a child is diagnosed with dyslalia, the Jaro method helps assess words where sound permutations were detected during reading as correctly pronounced, highlighting the identified problematic words in a different color in the output of the child's test results.

The results presented in Table 6 demonstrate that the Jaro method is capable of detecting sound permutations characteristic of children with dyslalia, significantly expediting text comparisons and enabling the identification of problematic words for the child, which constitutes a key advantage of this algorithm.

Table 6

Results of a comparison of the Loewenstein and Jaraud methods for a group of users who have problems with dyslalia

Text	Algorithm	Average Comparison Accuracy	Mean Time to Compare Two Texts When Dyslalia Is Detected	Sensitivity to permutations
TEXT A (60 words)	Levenshtein	91,3 %	1.47 s	-
	Jaro	92,5 %	1.2 s	93%
TEXT B (89 words)	Levenshtein	90,15 %	1.75 s	-
	Jaro	90,7 %	1.6 s	92,4%
TEXT C (135 words)	Levenshtein	86,3 %	2.1 s	-
	Jaro	85,9 %	1.95 s	91,8%

For the second experiment, it is shown that the Jaro algorithm also has a shorter comparison time for large text arrays (by 7%). The total time of operation of methods for comparing texts when dyslalia is detected in a child is due to the fact that in this violation variants of text arrays are not generated, but only possible permutations of sounds are evaluated.

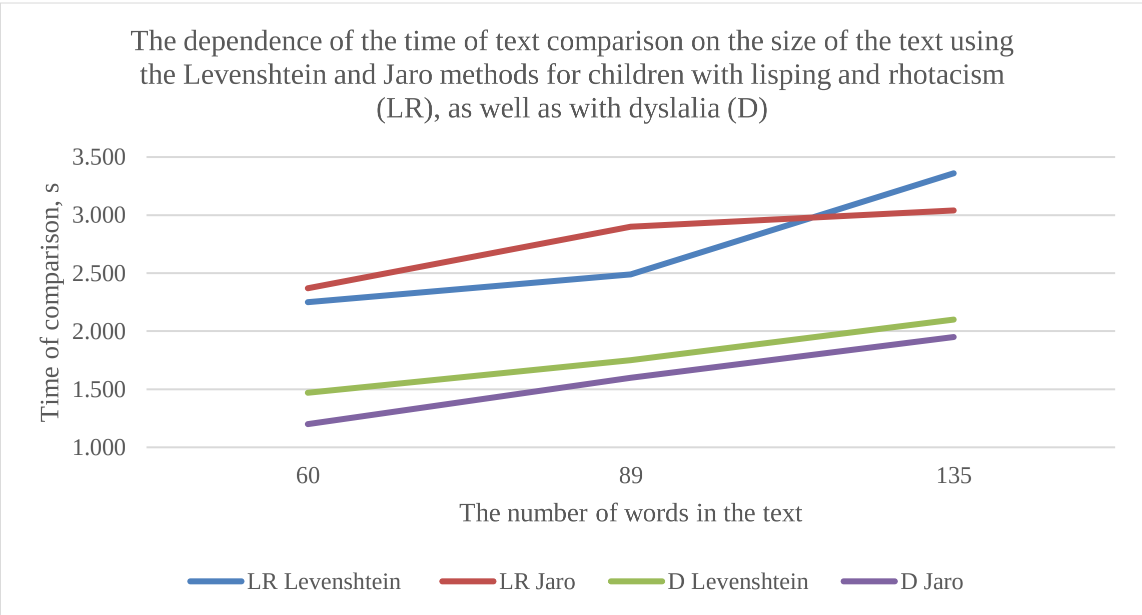

Figure 7. Analysis of the results obtained

Therefore, as a general conclusion, it can be stated that for short texts in the absence of speech disorders in children, or in cases of detected lisping and rhotacism, the Levenshtein method is preferable (Figure 7). In all other instances, the method of comparing texts by Jaro should be chosen.

The developed Clean Reading system has a simple interface understandable for children (Figure 8). The first step is to identify the child's speech defects by reading control words or texts. To start working with the child's voice, it is necessary to provide access to the microphone, after which the user can start direct testing. The user can choose the next action in the left menu (Figure 8a). The progress of recording the child's voice is displayed in the form of a microphone and a timeline, which is calculated for 1 minute of text reading (Figure 8b).

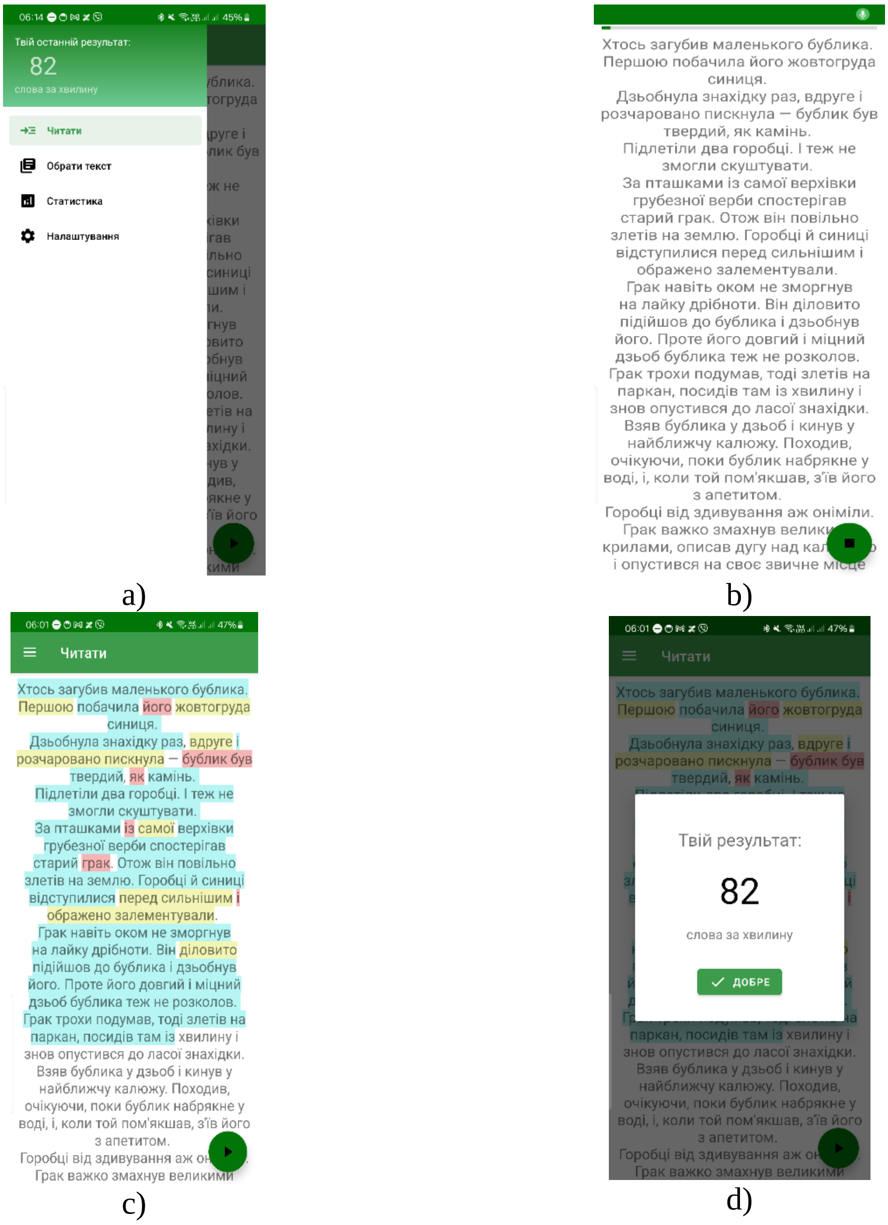

Figure 8. User interface: a) next *action selection menu, b) the process of testing a child's voice recording, c) analysis of the correctness of the words read, d) the number of words read per minute*

The control words are generated differently each time. They are specific, recommended to identify problems with the pronunciation of certain sounds, which are often inherent in children of younger grades. For example, the sounds [p], [p'], [l], [l'], [sh], [s], [s'], etc.

The last step is the analysis of the read text and the calculation of the result – the number of correctly and incorrectly read words (Figure 8c).

The results of the test are stored for the formation of statistics and the progress of improving the technique and reading speed (Figure 8d).

4. CONCLUSIONS AND PROSPECTS FOR FURTHER RESEARCH

Within the scope of this work, a simulator for an educational training tool was developed to monitor the technique and reading speed of elementary school students for a mobile digital device. The proposed educational game incorporates voice recognition technology with a high level of accuracy (up to 94.84%), identifying incorrectly read words and calculating the number of words read within a specified time frame.

The architecture of the proposed reading technique control system consists of the following modules:

– preparation module, which ensures the correct functioning of the system, regardless of the existing speech disorders;
– speech-to-text converter directly converts the voice of the child reading the text on the screen into text;
– the text comparison module is responsible for comparing the text read aloud with the text obtained in the conversion result.

In order to achieve the set objectives, an analysis of the requirements and principles for creating a mobile game for children was conducted, taking into account the age-specific characteristics of the target audience and potential speech impairments (articulation disorders, dyslalia, tachylalia, stuttering). The architecture of the proposed system was developed, services of Speech-to-Text (STT) were explored based on criteria such as accessibility, dependency on Internet connectivity, and computational resource requirements. Methods of STT for analyzing the accuracy of text read by a child were investigated by determining the similarity of strings through character-by-character comparison, specifically utilizing the Levenshtein and Jaro methods.

The evaluation of methods employed during the comparison of text arrays was conducted across user categories exhibiting diverse speech impairments. The first category comprised children experiencing lisping and rhotacism issues. The second category encompassed users with dyslexia, particularly concerning sound rearrangement within words during reading. Therefore, the analysis of research findings is segmented into two distinct conclusions:

– during the testing of children with lisping and rhotacism issues, it was observed that the Jaro algorithm exhibits slightly higher accuracy in comparing two texts (on average by 1.14%) and, concurrently, requires less time for comparison with large text arrays (by 39%);
– during the testing of children with dyslalia issues, particularly in terms of sound rearrangement in words during reading, it was also found that the Jaro algorithm exhibits a shorter comparison time for large-scale text arrays (by 7%). The reduced operational time of text comparison methods in detecting dyslalia in children is explained by the fact that this disorder does not generate variations of text arrays, but only evaluates possible sound permutations.

Therefore, as a general conclusion, it can be stated that for short texts in the absence of speech disorders in a child, or in cases of detecting lisping and rhotacism, the preferable choice is the Levenshtein method (Figure 7). In all other instances, the method of text comparison by Jaro should be selected.

The prospects for further research in this work involve further improvement of machine learning algorithms aimed at understanding different accents and dialects, not just the speaker's speech impairments. The personalization and adaptability of intelligent models will be expanded by incorporating multimodal data learning, leading to improved training modules, relevant feedback, and adjustment of exercise difficulty based on user performance.

The simulator's functionality is planned to be expanded with modules for mentoring control, collaborative learning, and social features, where peers can be involved in collaborative learning, group projects, and discussion forums.

REFERENCES (TRANSLATED AND TRANSLITERATED)

[1] L. Serice, “Prisms of Neuroscience: Frameworks for Thinking About Educational Gamification,” *AI Comput. Sci. and Robot. Technol.*, vol. 2, Mar. 2023, doi: 10.5772/acrt.13. (in English)
[2] S. Papadakis, “Tools for evaluating educational apps for young children: a systematic review of the literature,” *Interactive Technol. and Smart Educ*., vol. 18, no. 1, pp. 18–49, 2021, doi: 10.1108/ITSE-08-2020-0127. (in English)
[3] M. Sanromà-Giménez, J. L. Lázaro Cantabrana, M. Usart Rodríguez, M. Gisbert-Cervera, “Design and validation of an assessment tool for educational mobile applications used with autistic learners,” *J. of new approaches in educational research*, vol. 10, no. 1, pp. 101–121, 2021, doi: 10.7821/naer.2021.1.574. (in English)
[4] Z. Zainuddin, S. K. W. Chu, M. Shujahat, C. J. Perera, “The impact of gamification on learning and instruction: A systematic review of empirical evidence,” *Educ. Res. Rev.*, vol. 30, p. 100326, 2020, doi: 10.1016/j.edurev.2020.100326. (in English)
[5] A. Christopoulos, S. Mystakidis, “Gamification in education,” *Encyclopedia*, vol. 3, no. 4, pp. 1223–1243, 2023, doi: 10.3390/encyclopedia3040089. (in English)
[6] W. Oliveira, J. Hamari, L. Shi, A. M. Toda, L. Rodrigues, P. T. Palomino, S. Isotani, “Tailored gamification in education: A literature review and future agenda,” *Educ. and Inf. Technologies*, vol. 28, no. 1, pp. 373–406, Jan. 2023, doi: 10.1007/s10639-022-11122-4. (in English)
[7] N. Limantara, N. Meyliana, F. L. Gaol, H. Prabowo, “Designing Gamified Learning Management Systems for Higher Education,” *Int. J. of Inf. and Educ. Technol.*, vol. 13, no. 1, pp. 25–32, Jan. 2023, doi: 10.18178/ijiet.2023.13.1.1776. (in English)
[8] J. Majuri, J. Koivisto, J. Hamari. “Gamification of Education and Learning: A Review of Empirical Literature,” in *Proc. 2nd Int. GamiFIN Conf.*, GamiFIN, 2018, vol. 2186, pp. 11–19. (in English)
[9] M. Sailer, L. Homner, “The gamification of learning: A meta-analysis,” *Educational Psychology Review*, vol. 32, no. 1, pp. 77–112, Mar. 2020, doi: 10.1007/s10648-019-09498-w. (in English)
[10] R. Jayakanthan, “Application of computer games in the field of education,” *The Electronic Library*, vol. 20, no. 2, pp. 98–102, Apr. 2002, doi: 10.1108/02640470210697471. (in English)
[11] R. E. Mayer, “Computer games in education,” *Annu. Rev. of Psychol.*, vol. 70, pp. 531–549, Jan. 2019, doi: 10.1146/annurev-psych-010418-102744. (in English)
[12] S. M. Ivanova, “The influence of computer games on forming of elements of logical thought for children of senior preschool age,” *Inf. Technologies and Learn. Tools*, vol. 2, no. 1, Sep. 2010, doi: 10.33407/itlt.v2i1.279. (in Ukrainian)
[13] M. Yaremiy, O. Galyuka, “Gamification as a means of forming the communicative competence of elementary school students,” *Bull. of Sci. and Educ.*, vol. 17, no. 11, 2023. (in Ukrainian)
[14] Yu. Bludova, O. Ilyina, “Gamification of the educational process in the New Ukrainian School,” *Humanities science current issues*, vol. 27, no. 1, pp. 169–173, 2020, doi: 10.24919/2308-4863.1/27.203390. (in Ukrainian)
[15] I. Mykhailichenko, H. Ivashchenko, O. Barkovska, O. Liashenko, “Application of deep neural network for real-time voice command recognition,” in *Proc. 2022 IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, 2022, pp. 1–4, doi: 10.1109/KhPIWeek57572.2022.9916473. (in English)
[16] O. Barkovska, D. Pyvovarova, V. Kholiev, H. Ivashchenko, D. Rosinskiy, “Information Object Storage Model with Accelerated Text Processing Methods,” in *Proc. 5th Int. Conf. on Computational Linguistics and Intelligent Systems (COLINS 2021)* Lviv, Ukraine, 2021, pp. 286–299.
[17] S. King, J. Boyer, T. Bell, A. Estapa, “An Automated Virtual Reality Training System for Teacher-Student Interaction: A Randomized Controlled Trial,” *JMIR Serious Games*, vol. 10, no. 4, p. e41097, 2022, doi: 10.2196/41097. (in English)
[18] P. Salehi, S. Z. Hassan, M. Lammerse, S. S. Sabet, I. Riiser, R. K. Røed, M. S. Johnson, V. Thambawita, S. A. Hicks, M. Powell, “Synthesizing a Talking Child Avatar to Train Interviewers Working with Maltreated Children,” *Big Data and Cognitive Computing*, vol. 6, no. 2, p. 62, 2022, doi: 10.3390/bdcc6020062. (in English)
[19] Muthmainnah, P. M. Ibna Seraj, I. Oteir, “Playing with AI to Investigate Human-Computer Interaction Technology and Improving Critical Thinking Skills to Pursue 21st Century Age,” *Education Research International*, p. 6468995, 2022, doi: 10.1155/2022/6468995. (in English)
[20] D. N. Aini, S. Setiawan, S. Anam, R. Masyitoh, W. Ibad, “Artificial Intelligence Learning Platform for Promoting Speaking Competence: A Suggested Scaffolding for Autonomous Learning in Pesantren,” in *Proc. 4th Int. Conf. on English Language Teaching (ICON-ELT 2023)*, pp. 131–144, 2023, doi: 10.2991/978-2-38476-120-3_13. (in English)

[21] O. Barkovska, Y. Liapin, T. Muzyka, I. Ryndyk, P. Botnar, "Gaze Direction Monitoring Model in Computer System for Academic Performance Assessment," *Information Technologies and Learn. Tools*, vol. 99, no. 1, pp. 63–75, Feb. 2024, doi: 10.33407/itlt.v99i1.5503. (in English)

[22] Z. Wu, X. Chen, S. U. J. Lee, "A systematic literature review on Android-specific smells," *Journal of Syst. and Softw*, vol. 201, p. 111677, Jul. 2023, doi: 10.1016/j.jss.2023.111677. (in English)

[23] A. Adekotujo, A. Odumabo, A. Ademola, O. Aiyeniko, "A Comparative Study of Operating Systems: Case of Windows, UNIX, Linux, Mac, Android and iOS," *Int. J. of Comput. Appl.*, vol. 176, pp. 16–23, 2020, doi: 10.5120/ijca2020920494. (in English)

Text of the article was accepted by Editorial Team 06.03.2024

НАВЧАЛЬНИЙ ТРЕНАЖЕР З ФУНКЦІЄЮ КОНТРОЛЮ ТЕХНІКИ ТА ШВИДКОСТІ ЧИТАННЯ НА ОСНОВІ МЕТОДІВ ПЕРЕТВОРЕННЯ МОВИ В ТЕКСТ

Олеся Барковська
кандидат технічних наук, доцент, доцент кафедри Електронних обчислювальних машин
Харківський національний університет радіоелектроніки, м. Харків, Україна
ORCID ID 0000-0001-7496-4353
olesia.barkovska@nure.ua

Георгій Іващенко
кандидат технічних наук, доцент, доцент кафедри Електронних обчислювальних машин
Харківський національний університет радіоелектроніки, м. Харків, Україна
ORCID ID 0000-0003-1027-5262
heorhii.ivashchenko@nure.ua

Дмитро Росінський
старший викладач кафедри Електронних обчислювальних машин
Харківський національний університет радіоелектроніки, м. Харків, Україна
ORCID ID 0000-0002-0725-392X
dmytro.rosinskyi@nure.ua

Даніїл Захаров
магістрант кафедри Електронних обчислювальних машин
Харківський національний університет радіоелектроніки, м. Харків, Україна
ORCID ID 0009-0006-4676-6599
daniil.zakharov@nure.ua

Анотація. Роботу присвячено актуальному питанню – гейміфікації навчального процесу школярів молодших класів із застосуванням цифрових мобільних пристроїв на прикладі розробки симулятора навчального тренажера для контролю техніки та швидкості читання. Практична новизна дослідження полягає в можливості оптимізувати процес контролю техніки читання дітей з різними вадами мовлення, а саме – ліспінгом, ротацизмом та дислалією, що є невирішеною проблемою для існуючих комп'ютерних лінгвістичних моделей. Практична значущість полягає в тому, що використання інформаційно-комунікаційних технологій та гейміфікація процесу навчання мотивує до отримання знань та заохочує до навчання не тільки за рахунок ігрової стратегії, а й за рахунок того, що дитина отримує схвалення, можливість виконати завдання наново, удосконалюючи свої навички. Запропонований та протестований симулятор навчального тренажера для контролю техніки та швидкості читання школярів молодших класів для мобільного цифрового пристрою розпізнає голос дитини, із високою точністю (до 94,84%), виявляє неправильно прочитані слова та виконує підрахунок кількості прочитаних слів за відведений час. Архітектура запропонованої системи складається з наступних модулів: модуль підготовки, який забезпечує коректність функціонування системи не залежно від наявних мовленнєвих порушень; speech-to-text перетворювач безпосередньо конвертує голос дитини, яка читає текст на екрані гаджету, у текст; модуль порівняння текстів відповідає за порівняння тексту, прочитаного вголос, із текстом, отриманим у результаті конвертації. Результати тестування дітей з проблемами ліспінгу та ротацизму показали, що алгоритм Жаро має трохи вищу точність порівняння двох текстів (у середньому на 1.14%), і водночас більш короткий час порівняння для великих розмірів текстових масивів (на 39%). Результати тестування дітей з проблемою дислексії, а саме перестановка звуків у словах при читанні, також показали, що алгоритм Жаро має більш короткий час порівняння рядків для великих розмірів текстових масивів (на 7%). Більш короткий час роботи методів порівняння текстів при виявленні дислалії у дитини пояснюється тим, що при цьому порушенні не генеруються варіанти текстових масивів, а лише оцінюються можливі перестановки звуків.

Ключові слова: комп’ютерна гра; техніка читання; мобільні пристрої; редакційна відстань; STT.