Abstract
At present, modern information technologies are developing very rapidly. A lot of different software tools are being created to improve and simplify people’s lives. This is especially true in the context of distance learning. In the context of online learning, the traditional form of conducting dictations needs to be transformed. Therefore, it is possible to propose automation of the process of conducting and checking dictations using software. The paper analyzes scientific research and publications of the current state of language synthesis technologies and text similarity testing. With the help of C# and the language synthesis libraries of Microsoft, Google, Amazon, software has been developed that allows the user to listen and type text, and then automatically check it with the initial sample. The conducted testing for different types of errors showed the possibility and expediency of developing the system for conducting and checking dictations. Follow-up activities will focus on improving the effectiveness of the program. The use of such programs will help improve the organization of conducting and checking dictations during distance learning and self-training of students.

Keywords: digital education; dictation; spelling; language synthesizer; text comparison automation; syntax check; personalized learning.

INTRODUCTION

Significant rates of scientific, technical and social progress, various crisis phenomena that arise in the modern world, inevitably leave an impression on the education system and create additional difficulties in the formation and upbringing of the young generation. In addition, traditional pedagogical means of education, content and organization of the educational process are increasingly ineffective. Crisis phenomena in pedagogy arise due to the inconsistency of the pace and nature of social and pedagogical processes (Pinchuk, Sokolyuk, Burov, & Shyshkina, 2019).

The changes that are taking place require adjusting all processes to them, which creates new problems, the solution of which needs time. Thus, education systems must inevitably go through reforms. Today we live in the era of information technology, so their use is mandatory in the field of education. An important task is the integration of digital technologies in teaching and learning. The use of innovative technologies and digital education tools, the development of digital learning environments are very relevant (Pinchuk, Sokolyuk, Burov, & Shyshkina, 2019; Litvinova, Burov, & Semerikov, 2021).

The Ukrainian education system already has the experience of using distance education during the coronavirus pandemic. The use of this experience now, during a full-scale war, becomes especially relevant and very important. Many educational
institutions are forced to work remotely, which affects the quality of learning material by students. However, conducting online training makes it difficult to perform certain types of work that were everyday for the usual training format. Ways to improve the quality of education can be used: search for new forms and methods of education; widespread use of Internet resources and mobile applications; use of the latest digital content (Litvinova, Burov, & Semerikov, 2021; Piatykop, Pronina, Tymofieieva, & Paliî, 2022; Kozlovsky, & Kravtsov, 2018).

The transition to distance learning requires the direct use of information technologies and the creation of new learning tools. Teachers and students of education have to get used to the use of new forms of learning and knowledge control. In general, distance learning requires transferring to a virtual environment the main forms of learning, such as groups, courses, tests, libraries, virtual rooms and whiteboards. Currently, many specialized tools are being created for conducting distance learning, including the use of video conferences, shared boards, tools for testing, checking and monitoring knowledge, conducting virtual laboratory work, etc. (Litvinova, Burov, & Semerikov, 2021; Piatykop, Pronina, Tymofieieva, & Paliî, 2022; Kozlovsky, & Kravtsov, 2018).

One of the means of learning that requires a new form in a distance format is dictation. Dictation is a method of testing knowledge in which the student writes a text by ear, which is dictated by the teacher. Dictation is a teaching method for practice. Its essence for students is that words, sentences are perceived by ear in the recording. This is one of the most effective forms of work. It contributes to the formation of punctuation skills, the development of strong spelling skills, as well as the consolidation of knowledge of grammar, vocabulary, and phonetics.

Also, dictation teaches students how to concentrate, develops hearing, vision, memory and develops the ability to use spelling rules in practice. All this gives reason to recommend dictation as one of the main types of work during spelling learning.

There are several types of dictation, each of which has its own function for learning. Each of them is important in the process of learning a language.

In the past, the use of dictation was an integral part of any teaching and learning of a language, because they contribute to better memorization of words and spelling. But at this stage of life, due to external reasons, people began to switch to distance learning, which made it more difficult to carry out this type of work. Therefore, writing dictations almost ceased to exist, although it is a very effective form of work. Dictation remains a powerful language learning tool. First of all, dictation practices speaking competences in listening and writing, and also improves lexical and grammatical competence. At the same time, the formation of speech competence in reading also takes place, because the one who wrote the dictation checks its correctness (Bihych & Rusnak, 2023).

Considering the fact that a significant part of the modern generation are visual learners who prefer digital means (Bihych & Rusnak, 2023), educational portals for learning foreign languages are also rapidly developing, which also offer dictation among other tasks. To conduct dictations in an online format, it is possible to use a speech synthesizer that translates the written text into oral, that is, it reads the text instead of the teacher.

Thus, the use of tools for writing dictations will allow in the future to use all the advantages of this teaching method and lead to an increase in the literacy of students.
Therefore, the purpose of the work is digital transformation the organization of conducting and checking dictations during distance learning and for self-training of students. To do this, the problem of developing a training system for conducting and checking dictations is solved. This system is based on the use of language synthesis tools and methods for measuring the similarity of text documents to search for errors in the text of the dictation.

**Theoretical Fundamentals of the Research**

The use of natural language processing (NLP) is gaining popularity. This part of Information Technology, Artificial Intelligence and Linguistics deals with the study of natural language analysis and synthesis.

Today's use of NLP often involves solving tasks of search, selection of contextual online advertising, translation, sentiment analysis for marketing tasks, speech recognition and synthesis, chatbots and voice assistants (Onyshchenko, Daniiel, & Kameniev, 2020; Pronina & Piatykop, 2023).

The use of text-to-speech synthesizers (TTS) is gaining particular popularity. Such programs allow to translate written information (text) into oral information (sound). Thanks to the use of speech synthesizers, for example, people can learn new foreign words with the correct pronunciation, voice books, messages, documents, build various search queries without using the keyboard, and much more.

The author of the paper (Zadorizhna, 2020) draws attention to the fact that the expansion of the functions of virtual assistants allows the use of speech synthesizers in the educational field. And this, in turn, allows to reduce the load on both teachers and students. The author also emphasizes the need to create expressive synthetic speech to realize the possibility of tuning artificial voices.

The paper (Ryzhkyova, 2019) emphasizes the advantages of speech synthesis programs in education. There notes that their use makes it possible to optimize the learning process, in particular, in the conditions of multicultural communication in the Ukrainian educational space. Thus, it is proposed to voice scientific, technical and educational texts for foreign students of higher education, which will allow to overcome the linguistic barrier. Voiceover of dictation texts in lessons in educational institutions is also increasingly practiced.

There are many articles that discuss different methods of using the speech synthesizer and improving it.

In a study of voice synthesis (Tkachenko, Tkachenko, & Shcherbatiuk, 2020), the authors concluded that a significant number of existing synthesizers are available only for work on specific operating systems. They propose a mathematical justification and development of a platform-independent speech synthesizer.

Among the methods for speech synthesis, the following two are distinguished: concatenative and parametric synthesis. The first method involves selecting the necessary speech units from a database of available signals according to the context and combining them to produce speech. Such synthesis can be based on linear prediction of coefficients (LPC), which has been successfully used for single-word synthesis, or on PSOLA (Pitch Synchronous Overlap and Add), which pays more attention to prosody control and modification (Ning, He, Wu, Xing, & Zhang, 2019).
In the second method, the human voice process is considered as a simulation. In it, the letters of the language are the sources of signals for the excitation of the digital filter. "Vocal organ parametric synthesis, formant parametric synthesis, HMM-based speech synthesis (Hidden Markov Model), and deep neural network (DNN)-based speech" are often used here (Ning, He, Wu, Xing, & Zhang, 2019).

In the study (Vainer & Dušek, 2020), a system for improving speech synthesis is proposed, which is capable of rapid learning, high-quality synthesis and output of sound. The described model involves the synthesis of logarithmically scaled mel spectrograms using phonemes as input data. The authors note that the system provides significantly better voice output quality than strong baselines and involves a combination with the MelGAN vocoder. The authors used «the sum of MAE and structural similarity index (SSIM) losses for logarithmic mel spectrogram value regression and Huber loss for logarithmic duration prediction» (Vainer & Dušek, 2020).

Instead of using Mel-spectrograms for conditional auto-encoding in work (Raitio, Rasipuram, & Castellani, 2020) “prosodically meaningful acoustic features derived from the speech signal” are used, which allows learning “only perceptually relevant acoustic differences that contribute to prosody”. The authors predict “the acoustic characteristics directly from encoder outputs with a focused set of prosodic features” and use a "unified encoder-decoder architecture" that can be trained as a single model. The model also uses the spectral tilt as the fourth dimension of prosody, which allows creating different speech styles with sufficient quality.

In the article (Adam, 2020) the author draws attention to the synthesis based on deep learning. Using a deep neural network (DNN) for speech synthesis, it is possible to create a basis for predicting the acoustic characteristics of parameters. The proposed model is able to integrate the extraction and modeling of acoustic features without the problem of fragmentation.

Deep learning is a method of machine learning based on the structure of artificial neural networks. Such a structure “can be trained using supervised and unsupervised learning algorithms” (Striuk & Kondratenko, 2021).

In the study (Ubskii, Matveev, & Minker, 2019), the authors separately valuate “the model’s performance on matrix language words and embedded words”. Two data sets are used in the work: one language with code-mixing and the other without code-mixing. Code-mixing is the alternation of languages in conversation, often used in multilingual communities. After that, a monolingual model is trained for each data set and a bilingual model is trained on their combination.

And, for example, the authors of the study (Weiss, Skerry-Ryan, Battenberg, Mariooryad, & Kingma, 2021) propose the sequence-to-sequence network capable of synthesizing speech from text. Their Wave-Tacotron model is based on the Tacotron model, but uses “a normalizing flow in the autoregressive decoder loop”. Experiments conducted by the authors prove a significant speed of generation and quality of speech synthesis using the proposed model.

Thus, the level of development of digital technologies for solving the problem of speech reproduction is quite high. Therefore, existing technology for speech synthesis can be used. We need to choose the best software for playing Ukrainian text. To check the text of the dictation for the correct spelling, it was decided to use our own approach.
TECHNOLOGIES AND METHODS FOR AUTOMATING THE CONDUCT AND VERIFICATION OF DICTATIONS

The choice of speech synthesizer

There are many processes and phenomena, quantitative information for the characteristics of which is not available or is changing rapidly. In this case, methods of expert assessments are used, the essence of which is that the choice is based on the opinion of a specialist based on professional, scientific and practical experience.

To select a library that will voice the text for the student, it was decided to use the expert assessment method. This method involves grading by multiple experts. Since there are several experts, then it is necessary to generalize their assessments.

The experts assessed the parameters in a quantitative scale, then it is necessary to build a final group assessment. For this, it was decided to use the average value.

The essence of the method is that each expert evaluates the factors in the range from 1 to 10. For example, the beginning of the scale - 1 point - poor performance on this criterion / parameter. The upper limit of the scale - 10 points corresponds to good work on the selected criterion / parameter. Using a similar scale, the expert must evaluate each parameter with a numerical value \( P_{ij} \) within the limits of the single scale used by him. The next step is to calculate the average value for all experts \( P_i \).

\[
P_i = \frac{\sum_{j=1}^{N} P_{ij}}{N}
\]

where \( N \) – number of experts, \( P_{ij} \) – a numerical value within the single scale used by him.

The values \( P_i \) can be used to determine the significance coefficients and evaluate the quality indicators.

To determine the weight coefficient, the observed condition \( \sum_{i=1}^{n} M_i = 1 \), we obtain the value by the formula:

\[
M_i = \frac{P_i}{\sum_{i=1}^{n} P_i}
\]

It was decided to use the adjustment of the values of the experts' assessments in order to make them more convenient for comparison.

Next, we calculate the coefficient \( F_{ij} \), and with its help we find the comparison numbers for each library \( K_j \)

\[
F_{ij} = P_{ij} \times M_i, \quad (3)
\]

\[
K_j = \sum_{i=1}^{n} P_j \quad (4)
\]

Further, the optimal resource for voicing dictation texts was determined by the ranking method.

Three language synthesizer libraries were selected for evaluation. Such as: Microsoft; Google; Amazon.

The evaluation result for the Microsoft Language Synthesis Library is shown in Table 1, Google in Table 2, and Amazon in Table 3, where:

- "E1, E2, E3, E4, E5, E6" – experts, that is, people who evaluated according to the criteria provided;
- "Parameter" – the criteria by which libraries will be compared (quality of pronunciation / comprehensibility, voice, correctness of stress, intonation, correctness of voicing punctuation marks);
- "$P_i$" – the average value of the results of the survey of experts for each parameter calculated using formula (1).

The panelists inserted scores from 1 to 10 for each library for each criterion. After that, a performance score will be calculated to identify the best library.

Table 1. Result of peer review for the Microsoft Language Synthesis Library

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameter</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
<th>E5</th>
<th>E6</th>
<th>$P_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Quality of pronunciation / comprehensibility</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Voice</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Correctness of accents</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>Intonation</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Pronunciation of punctuation marks</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>7</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 2. The result of the peer review of the Google Language Synthesis Library

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameter</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
<th>E5</th>
<th>E6</th>
<th>$P_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Quality of pronunciation / comprehensibility</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>Voice</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>Correctness of accents</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>Intonation</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>Pronunciation of punctuation marks</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 3. Peer review result for Amazon Language Synthesis Library

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameter</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
<th>E5</th>
<th>E6</th>
<th>$P_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Quality of pronunciation / comprehensibility</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>Voice</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>Correctness of accents</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>Intonation</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>Pronunciation of punctuation marks</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 4 shows the overall result of the evaluation of experts on the criteria for each language synthesizer library. Also, in table 4, the column "Weight" is added – the weight of the parameter, depending on the degree of importance (where the sum is 1). Can be calculated using formula 2.8, or selected independently based on the needs of the use of parameters / criteria for evaluating the libraries present.

Note that the quality of pronunciation / comprehensibility is the main parameter, so we will give it more points, then the correctness of voicing punctuation marks, as well as the voice, the correctness of stress, intonation, which are no longer so important when writing a dictation.

Let’s add 4 rows to the table "Sum" – the number obtained as a result of compiling the "weight" for each library.
Next, we calculate the coefficient $F_{ij}$, and with its help we find the comparison numbers for each library $K_j$, that is, the sum according to formulas (3), (4).

The result of multiplying expert scores by "weight" is shown in Table 5.

Table 4. Overall result according to experts

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameter</th>
<th>Microsoft</th>
<th>Google</th>
<th>Amazon</th>
<th>Weight ($M_i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Quality of pronunciation / comprehensibility</td>
<td>2</td>
<td>9</td>
<td>9</td>
<td>0.4</td>
</tr>
<tr>
<td>2</td>
<td>Voice</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>Correctness of accents</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>Intonation</td>
<td>3</td>
<td>9</td>
<td>8</td>
<td>0.1</td>
</tr>
<tr>
<td>5</td>
<td>Pronunciation of punctuation marks</td>
<td>9</td>
<td>5</td>
<td>5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Table 5. Full result of the peer review

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameter</th>
<th>Microsoft</th>
<th>Google</th>
<th>Amazon</th>
<th>Weight ($M_i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Quality of pronunciation / comprehensibility</td>
<td>0.8</td>
<td>3.6</td>
<td>3.6</td>
<td>0.4</td>
</tr>
<tr>
<td>2</td>
<td>Voice</td>
<td>0.5</td>
<td>0.9</td>
<td>0.9</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>Correctness of accents</td>
<td>0.5</td>
<td>0.9</td>
<td>0.9</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>Intonation</td>
<td>0.3</td>
<td>0.9</td>
<td>0.8</td>
<td>0.1</td>
</tr>
<tr>
<td>5</td>
<td>Pronunciation of punctuation marks</td>
<td>2.7</td>
<td>1.5</td>
<td>1.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

| Sum ($\sum K_j$) | 4.8 | 7.8 | 7.7 | 1 |

Thus, according to experts, with the help of calculations, it can be concluded that it is better to use the Google library in work, since it scored 7.8 points, followed by Amazon, equal to 7.7 points, after Microsoft, with 4.8.

Method for assessing the similarity of two texts

The assessment of the proximity of two texts $t_1$ and $t_2$ is carried out by evaluating its objects – sentences. Moreover, $t_1$ is the reference text with which $t_2$ is compared.

There are two models of application of texts $S_1$ and $S_2$ respectively belonging to texts $t_1$ and $t_2$. Suggestions are taken to evaluate these texts. $s_1^1 \in S_1$ and $s_2^2 \in S_2$. The evaluation is carried out on pairs of sentences from each text, taking into account the number of the sentence in order. The comparison is carried out taking into account the occurrences of the words $w_j$ in sentences. The general comparison of $S_1$ and $S_2$ occurs by comparing the occurrences of the words $w_j$ in the texts.

The general text comparison model is presented in the following form:

$$ T = \{(w^1, w^2) \in W, (w^1 \in s_1^1), (w^2 \in s_2^2), (S_1 \subset t_1), (S_2 \subset t_2)\} $$

(5)

where $(w^1, w^2) \in W$ – words from two compared texts that are included in the text dictionary; $W$ – the dictionary of the text, namely the set of all words of the reference text.

A similarity measure is used to evaluate the comparison of two texts. In case of mismatch of the forms of the word, the proximity of the words will not be equal $\delta(w_j^1) \neq \delta(w_j^2)$. 

Since the main task of comparing texts is taking into account the comparison of the text when writing test papers, namely dictations. It is necessary to take into account the degree of importance of the word, as well as typos (misprints) in the text.

The degree of importance of a word is calculated based on the TF-IDF methodology (Kim & Gil, 2019), namely the calculation of a statistical measure of the importance of a word in a document.

Some words in the text occur more frequently and, on the one hand, they may not be important when checking, since they can be elementary words or conjunctions, and on the other hand, the more often a word is used, the less likely it should be to make a mistake in this word.

\[ tf(w, t) = \frac{n_k}{\sum n_k} \]  
\[ idf(w, T) = \log_{10} \left( \frac{|T|}{|\{t_i \in T, w \in t_i\}|} \right) \]

where \( tf(w, t) \) – the frequency of the word in the text; \( n_k \) is the number of occurrences of the word \( w \) in the text \( t \); \( \sum n_k \) is the total number of words in the text.

\[ \sum \] where \( \sum \) – the number of texts in the collection of documents; \( |\{t_i \in T, w \in t_i\}| \) – the number of texts from the collection \( T \) in which the word \( w \) occurs, and \( n_w \neq 0 \).

The general formula for calculating the degree of importance of a word is presented in the following form.

\[ v(w_j) = tf(w, t) \times idf(w, T) \]  

Estimating the degree of importance of a word is necessary to understand the level of student errors. You can also track error statistics by evaluating the degree of importance of a word. If it’s a commonly used word, then the error matters more.

In turn, for a typo or incorrect spelling of a word, it is necessary to provide a fine \( f(w^1, w^2) \)

\[ f(w^1, w^2) = \frac{l_k}{\sum l_k} \]

where \( l_k \) is the number of correctly spelled letters in a word; \( \sum l_k \) – the total number of letters in the word.

\[ f(w^1, w^2) = \begin{cases} 1, & \text{if the words are spelled correctly} \\ 0 < f(w^1, w^2) < 1, & \text{if the words are misspelled} \end{cases} \]

In addition, to assess the proximity between words, it was decided to also use the distance between words, similar to how the distance between points is calculated in the space of points. To do this, you need to set the rules by which it will be calculated. The greater this distance, the less similar the words and vice versa. To do this, the concept of a single-character operation is introduced. There are only three of them: insert - the student added a letter that should not be there; deletion – the student did not write the letter that should be there; substitution - the student misspelled a letter.

Thus, for each single-character operation, there is a value by which the similarity of two words can be compared. The Levenshtein distance (Ouarda, Malika, & Brahim,
between two words is defined as the minimum number of single-character operations required to compare two words. It was decided to assign a weight to each operation, since typos (mispellings) have less weight than misspelled words. Replacement – 0.6, removal – 0.5, insertion – 0.3.

It is best if the assignment of the weight of each operation is adjusted for the letter. Since if the student forgot about doubling the letters at the root of the word, this is a spelling mistake and the weight should be higher than if the student forgot to add the last letter of the word.

This method was chosen because when checking dictations, different spellings of words are possible, according to the number of letters, and this method works qualitatively with different line lengths. Unlike the metric for calculating the Hamming distance (Ouarda, Malika, & Brahim, 2023), where if the student skips a letter or writes an extra one and the length of the words for comparison is different, this method will not work. There is also a modification of the method for calculating the Levenshtein distance – the Damerau-Levenshtein distance (Ouarda, Malika, & Brahim, 2023). This method was not considered because the difference from the classical one in adding a transposition is the replacement of two characters in places. Due to the specifics of the task - checking dictations, it was decided that this method is not suitable. Since the permutation of characters is a variant of the error and is considered as a replacement of letters. The remaining methods for the task at hand are redundant and not suitable.

The system for conducting and checking dictations

To implement the prototype of the training system for conducting and checking dictations, the following conditions were determined:
- the need to work with a file containing at least 500 characters;
- the program must have a text-to-speech process for dictation;
- it is necessary to have such opportunities as stopping and turning on text playback;
- the ability to select various libraries for the synthesis of the language, to select a convenient one for a particular person;
- the ability to check the text, compare with the original and display them on the screen;
- conclusion of the analysis of the results of writing a dictation, such as “the number of words written in total”, “the number of punctuation marks”, “the number of errors”, as well as highlighting errors in the text itself.

To develop the program, libraries were used to create the user interface Win Forms (Windows Forms) and .NET Framework, and the module itself was written in the high-level C# programming language.

An example of a program window with the result is shown in Figure 1.
RESEARCH RESULTS

When comparing the results of the dictation with the original, the following types of errors were considered:

- detection of spelling errors (works well; highlights the word with an error in red in the original text, to draw attention to the correct spelling, and not taking into account the error; displays the number of errors in the output); the only negative is that it does not show the number of errors in one word, but only that there is an error in one word;
- skipping a punctuation mark (works well; all characters are highlighted in yellow in both texts so that the student can see for himself where his mistake is; in the output it displays fewer words and punctuation marks than necessary);
- extra punctuation marks (works well; all characters are highlighted in yellow in both texts so that the student can see for himself where his mistake is; in the output it displays more words and punctuation marks than necessary);
- writing all words in lowercase letters (works well; highlights only words that require capitalization, considers this an error);
- writing all words in capital letters (works well; highlights all words, considering this an error, because it is not clear whether the student correctly understood the spelling of words with capital letters or not);
- use of unnecessary spaces (works well; does not pay attention to it: it does not matter how many of them, because this does not apply to spelling and punctuation errors of the dictation);
- the use of special characters (it is impossible to determine the correctness of the work, since from the point of view of the program, two words are compared, and when the word includes an unnecessary letter or sign, this is an error; if the special character is separated from the word, then this is not perceived as an error; the output displays more words).
Checking defects in practice in the calculated form is given in Table 6.

Table 6. Text analysis example

<table>
<thead>
<tr>
<th>No.</th>
<th>Possible defects</th>
<th>Input data</th>
<th>Initial data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Spelling errors (1 per word throughout the text)</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>Spelling errors (several in 1 word throughout the text)</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>Skipping a punctuation mark</td>
<td>6</td>
<td>Distributed signs: 5 out of 11</td>
</tr>
<tr>
<td>4</td>
<td>Extra punctuation marks</td>
<td>15</td>
<td>Distributed signs: 15 out of 11</td>
</tr>
<tr>
<td>5</td>
<td>Skipping the desired punctuation mark and entering the extra</td>
<td>Missed 6, added</td>
<td>Distributed signs: 9 out of 11</td>
</tr>
<tr>
<td>6</td>
<td>Writing all words in lowercase letters</td>
<td>All words are given (78)</td>
<td>4 mistakes in words where a capital letter is required</td>
</tr>
<tr>
<td>7</td>
<td>Capitalize all words</td>
<td>All words are given (78)</td>
<td>61 errors (highlights almost all words)</td>
</tr>
<tr>
<td>8</td>
<td>Using unnecessary spaces</td>
<td>Added 9</td>
<td>Does not perceive as an error (0)</td>
</tr>
<tr>
<td>9</td>
<td>Space spaces (not gluing words together)</td>
<td>7</td>
<td>No errors, word count 78 out of 78</td>
</tr>
<tr>
<td>10</td>
<td>Space spaces (gluing two words together)</td>
<td>7</td>
<td>Number of words 71 out of 78 (perceived as 1 word), errors 14 (per word)</td>
</tr>
<tr>
<td>11</td>
<td>Space spaces (gluing multiple words together)</td>
<td>Missing 5 spaces (gluing into 1 word)</td>
<td>Number of words 73 out of 78 (perceived as 1 word), errors 5 (per word)</td>
</tr>
<tr>
<td>12</td>
<td>Use of special characters (not gluing with the word)</td>
<td>2</td>
<td>Number of words 80 out of 78, errors 0</td>
</tr>
<tr>
<td>13</td>
<td>Using special characters (gluing with a word)</td>
<td>2</td>
<td>Number of words 78 out of 78, errors 2</td>
</tr>
</tbody>
</table>

CONCLUSIONS

The education remains extremely relevant and important in today's world. Rapid progress in science and technology requires people to be willing to constantly learn and develop their skills, including learning new languages. Previously, dictations were an integral part of any language learning and learning for better memorization of words and spelling. But at this stage of life, due to various external reasons, education is often forced to be distance learning. Under these conditions, it became more difficult to carry out this type of work, so the writing of dictations almost ceased to exist. Although this is a very effective form of work. The main advantage of dictation as a method of testing knowledge is that it allows you to evaluate not only the knowledge of the student, but also his ability to correctly use the acquired knowledge in practice. Also, this method can help to identify mistakes that a student makes when writing a text.

The development of information technology makes it possible to create software for automating the process of conducting dictations. As a result of the study, using the method of experts, the speech synthesizer libraries of Microsoft, Google, Amazon were evaluated.
A prototype of a program has been developed that performs the pronunciation of the text and the subsequent verification of the text typed by the user with the original one. The conducted testing for different types of errors showed the possibility and expediency of developing a system for conducting and checking dictations. Follow-up actions will be aimed at improving the effectiveness of the program. The use of such programs will help improve the organization of conducting and checking dictations in distance learning and self-training of students.

REFERENCES


About the authors:

**Olena Piatykop**, State Higher Educational Institution "Pryazovskyi State Technical University", Dnipro, Ukraine. ORCID: https://orcid.org/0000-0002-7731-3051. piatykop_o_ye@pstu.edu

**Olha Pronina**, State Higher Educational Institution "Pryazovskyi State Technical University", Dnipro, Ukraine. ORCID: https://orcid.org/0000-0001-7085-8027. pronina.lelka@gmail.com

**Lyudmila Kotykhova**, State Higher Educational Institution "Pryazovskyi State Technical University", Dnipro, Ukraine. ORCID: https://orcid.org/0009-0006-5008-622X. kotsykhova_l_d@pstu.edu