Reinforcement Learning Algorithms to Model English Pronunciation

Yan Zhengkun*, Liu Qian
School of Language and Culture
Southwest University of Finance and Economics
Chengdu, 610000, Sichuan
China
loobcr3753143618@163.com

ABSTRACT: *With the development of artificial intelligence technology, computer-aided pronunciation training systems are gradually receiving attention. However, existing systems have some issues, such as inaccurate judgments and untimely feedback. We propose an improved solution based on deep learning and reinforcement learning to address these issues. This scheme first uses deep learning algorithms to model English pronunciation. Specifically, we used models such as Convolutional Neural Networks (CNN) and Long Short Term Memory Networks (LSTM) to extract and classify features of pronunciation audio. At the same time, we also used reinforcement learning algorithms to automatically adjust training difficulty and feedback methods based on students' pronunciation to provide more personalized and accurate training effects. Through experimental verification, we found that the improved scheme based on deep learning and reinforcement learning can significantly improve the accuracy and real-time performance of the CAPT system in English pronunciation judgment. This scheme has higher learning efficiency and lower learning costs than traditional pronunciation training systems. In addition, students also showed higher satisfaction and stronger learning motivation towards the program.*

Keywords: CAPT System, English Pronunciation, Judgments, Improvement

Received: 19 March 2023, Revised 9 June 2023, Accepted 17 June 2023

DOI: 10.6025/jcl/2023/14/3/73-81

1. Introduction

With the gradual maturity of computer technology, the information processing model represented by intelligent systems has moved towards a larger display platform [1]. In the current CAPT system, the computer pronunciation training software has pronunciation judgment errors in the application process, which not only causes some misleading to users but also reduces the trust of CAPT system [2]. Based on the above problems, this paper will improve the CAPT system based on the basic algorithm, starting from the collection of English pronunciation, then constructing the acoustic model for the syllable entry point and using the neural network model for the systematic reconstruction of the acoustic model [3].

During the correction process, the error analysis is firstly carried out based on the CAPT evaluation results, the homophonic errors in phonetic classification are classified, the HMM model is used to classify them overall, and the discernible pronunciation options are summarized [4]. Then the neural network model is introduced, and the acoustic processing algorithm is introduced step by step. Based on this background, it is easy to input and obtain the above algorithm's result, and a formal correction plan [5] is proposed. After the above layers of screening, it will fundamentally compensate for the error output of the CAPT model in the previous model, thereby improving the system's error correction capability, root out to solve the error output, the overall performance of the system is optimized [6]

2. State of the Art

CAPT system is a system design model first proposed by American scientist Barbour in the 1990s. It aims to use computer technology to provide efficient English pronunciation learning error correction purposes [7]. In this century, the United States held more than ten meetings on measures of CAPT system research and improvement and set up and improved the system algorithm in stages. Over more than ten years of constant revision, it has been widely applied in the United States and received highly positive feedback. However, due to the homophonic voice pronunciation in our country and the pronunciation of English words, error correction errors often occur in systematic error correction [8].

Compared with other countries, there are still some problematic language differences in our country's development of CAPT system. Due to the limitation of domestic computer technology and the final modification of the system in the United States, this problem needs to be adequately addressed [9]. Therefore, based on the basic algorithm, we apply the neural network model and the HMM model to correct the decision-making algorithm, input the homophonic part into the neural network model in the form of the data packet, construct the node by using the neural network, part of the filter to delete, the special algorithm is used to deal with the special situation, and finally, the whole optimization is carried out [10].

3. Methodology

3.1. Improved CAPT Network Algorithm Building

In the CAPT system, the collection of error judgment factors is essential to the system error correction analysis. This paper uses the neural network model to integrate the retrieval process and among the CAPT system, which led to many false decision factor data architectures. Then, according to the indication standard of each data obtained from the algorithm and the requirements of the specification, the information data is output for the homophone word unit information, and the model receives the data for digest processing and finally completes the construction of the entire acquisition model. In this process, the algorithm of the neural network model is adopted. First, the number of data that led to the error-determine factor in the CAPT system is obtained through the following calculation, the title of the CAPT system is encoded, and the specific coding is used to classify the unit data of homophonic words under the jurisdiction class, the neural network model is used to organize the system to provide preconditions for further work automatically. Based on this CAPT system which leads to the wrong decision factor coding, fitting the unit dataÿwe consider ∂, t as branches of the homophone word coefficient, then C is the number of shunt coefficients, T is the number of categories of cavity numbers, r is the number of homophonic words, and represents the number of homophonic words in the CAPT system, the direction of the value, through the following formula checks, come to homonymic word data classified integration program.

$$\frac{\partial}{\partial \chi}\left(\lambda \frac{\partial T}{\partial \chi}\right)+\frac{\partial}{\partial y}\left(\lambda \frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial z}\left(\lambda \frac{\partial T}{\partial z}\right)+Q=pc\frac{\partial T}{\partial t} \qquad (1)$$

In this formula, the main consideration is that homophone words can be transmitted to homophone word units at the maximum flow rate after the integration of the homophonic words is completed. A represents the number of each independent algorithm in the CAPT system that leads to the wrong decision factor. As these algorithms all participate in the overall load, the overall load capacity will be enhanced. In addition, Q and Z represent the expressiveness of its error correction design entry nodes, respectively, and v represents the error correction design node coefficients. And once a node fails, other nodes will be supplemented accordingly, that is, the description of our formula. When, $\partial(i) = l(i)$ the following formula will appear:

$$I[Q_1 + (v_0, v_5) + Q_2] = min\{ I(Q_1), Z_1 - f(v_5, v_6), I(Q_2)\} \qquad (2)$$

Finally, the final solution of the neural network model is determined after reading its optimal node coefficient. In the following formula, the error correction design node coefficient, *K, P* represents the deviation coefficient, and node 0 represents the paralysis of the network of the homophone word unit. The situation will not happen in general. Therefore, we set the value of ∂i to be greater than the value of ρi. May wish to assume that the ratio of ∂i and ρi, the approximate golden ratio, that $\partial i{:}\rho i$ = 0.3821: 0.6182. Due to the floating-point operation having a certain load, take the approximate ratio of 0.4: 0.6. That is, the value of is 0.6 and the value of is 0.4. Homophone word unit weight function expression can be expressed as:

$$\omega(\chi i) = \sum_{j-1}^{k} \partial i^2 + \sum_{i-1}^{p} \rho i^2 \qquad (3)$$

To ensure the regular operation of the algorithm: in this paper, the dynamic BEY language is used to judge the structure of the CAPT system, and the resources of the homophone word unit based on the error judgment factor are played according to the result of the judgment. Finally, the filtering result of the algorithm is uploaded into the processing part of the algorithm through the packing form, and the algorithm filters the information of the homophonic word unit, and finally, the integrated processing is performed, and the specific process is as shown in the following figure

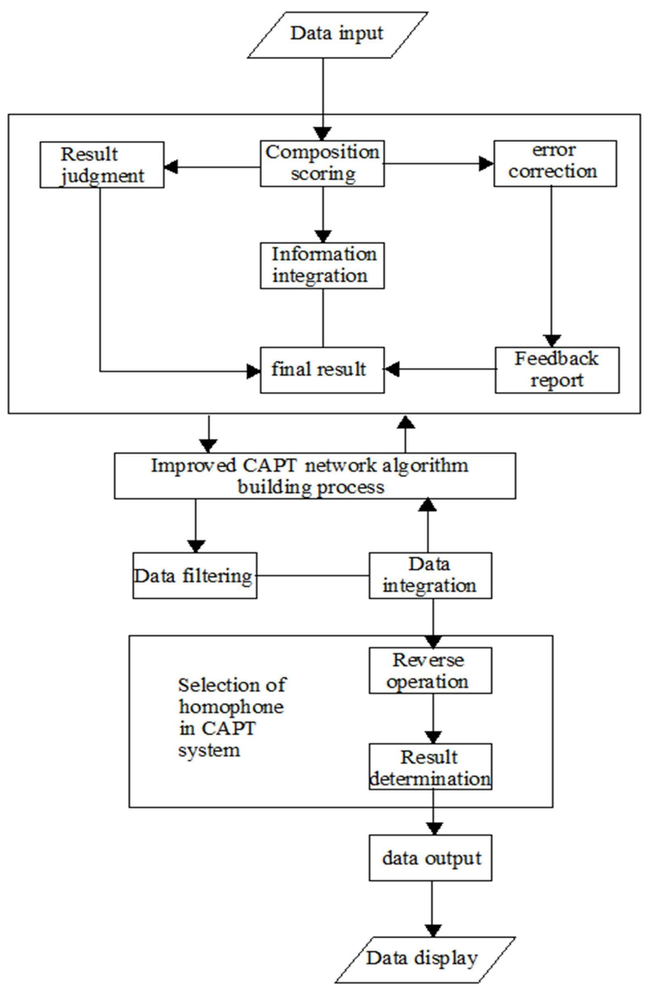

Figure 1. Improved CAPT network algorithm building process

3.2. Neural Network Model Overlap Algorithm

In the CAPT system, which leads to the establishment of the wrong decision factor model, the next one is to consider the CAPT system which leads to the wrong decision factor model and neural network word non-linear algorithm for technical integration. For the construction of the whole algorithm, this article divides it into four parts. Including the CAPT system, which led to the terminal error signal to determine the harmonic word collection, that is, in the neural network model based on the calculation model, from each system error analysis factor extracted from each integrated neural network model calculation model information data, and Basic homophonic word unit data for each phase. Among these are the information provided by the compendium of technologies that led to the misjudgment of the total CAPT system. Secondly, we deal with the data; the primary need to be completed is the information from different homophonic word unit curve integration and processing, effective processing by the algorithm, reached a consensus in all aspects to achieve the smooth operation of each algorithm. The operation of the algorithm has provided protection. Finally, it is imported into each system error correction analysis page, and the neural network model calculation models and information collected by each algorithm, as well as the components of the CAPT system that cause the error determination factors, are displayed on the overall algorithm results.

Error causing factors in CAPT system			**Computational model of neural network model**		
Factor	**Accuracy rate**	**Resonant adjustment**	**efficiency**	**Overall effect**	**Remarks**
Integrating and processing the information from different homophonic word unit curves	data packet1	0.55	0.86	0.74	
	data packet2	0.23	0.35	0.63	
	data packet3	0.52	0.51	0.82	
	data packet4	0.53	0.46	0.32	
	data packet5	0.65	0.36	0.78	
	data packet6	0.75	0.42	0.87	
Traditional inverse algorithm	data packet7	0.51	0.61	0.73	Consider stability
	data packet8	0.95	0.89	0.92	

Table 1. Component of Error Causing Factors in CAPT System

As shown in Table 1, in the words of dynamic algorithm of homophonic unit integration phase, we use *XY* model to represent the total data algorithm, the standard data model by *R* on behalf of the homophonic words, the algorithm in the information integration stage adds a set of accurate, objective model. In order to ensure the accuracy of the results, we need to reduce the algorithm error. The design of this link first through the two parameters to build a corresponding probability calculation, assuming $F(X, Y)$, input from the corresponding training algorithm, to get an optimal result Y. The formula model is as follows:

$$\begin{gathered}(X_i, Y_i), i = 1,2\ldots\ldots, \text{n}, X_i \in R^m, Y_i \in Y \\ (X_{i-1}, Y_{i-1}), i = 0,1\ldots\ldots, \text{n-1}, X_i \in R^m, Y_i \in Y \\ (X_{i-2}, Y_{i-2}), i = -1,0\ldots\ldots, \text{n-2}, X_i \in R^m, Y_i \in Y\end{gathered} \tag{4}$$

Using this formula is used to represent the optimal coefficient and F represents the standard coefficient. The minimum error risk can be obtained, and the error can be reduced to a minimum.

$$R(\omega)=\int L(Y,F(X,\omega))dF(X,Y)$$

$$R_l(\omega)=\frac{Y^{0.4}-1}{X^{0.4}-4} \tag{5}$$

I use the formula to make the corresponding prediction function, where *Remp*(ω) represents the set of predictors, *L* represents the optimal coefficient, represents the longitudinal coefficient, *f* (*Xiω*) use these data to establish the corresponding functional relationship, and gets the corresponding difference sequence. If you encounter the corresponding classification problem, we can take a collection function method, assuming the value *R* in the region (-1,1), the optimal function as follows:

$$R_{emp}(\omega)=\frac{1}{n}\sum_{i=1}^{n}L(Y_i,f(X_i\omega)) \tag{6}$$

After the prediction of the optimal numerical value, we aim at the probability of the output between the assumed algorithm *R* and the homophonic word unit algorithm *Q*. Finally, according to a series of optimal numerical values, the corresponding joint probability is calculated. From the formula (3), the joint probability can be calculated to minimize the expected risk. Since the joint probability is unknown, *R* is unable to compute the minimization directly. However, by using the large number theorem, according to the known training sample set, the arithmetical average can be used to ensure the accuracy of the scoring results. Each factor input algorithm flow as shown in figure 2.

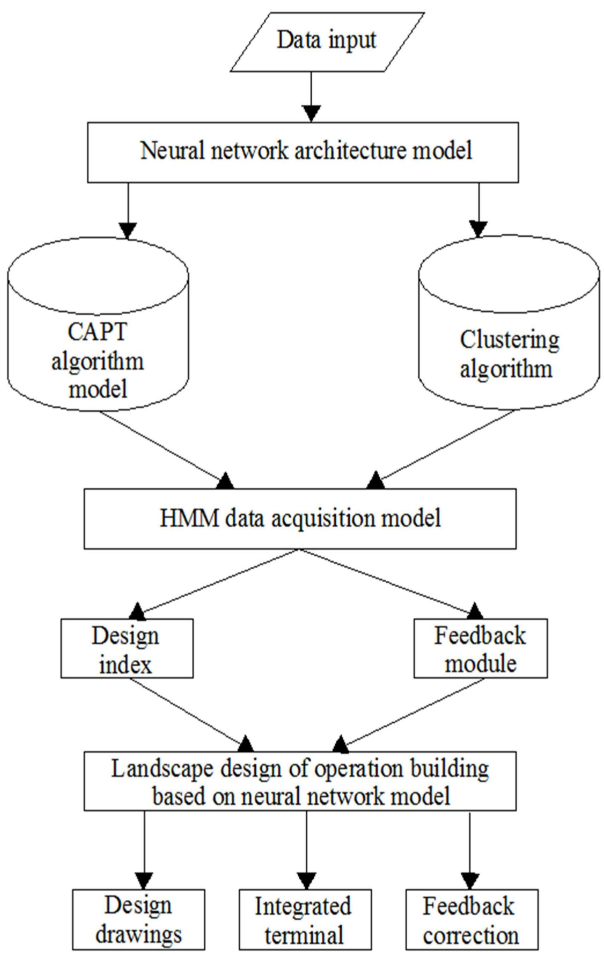

Figure 2. Cause Frame for Error Determination in CAPT System

Second, through the calculation and integration of information data, the design process will focus on CAPT system, which led to the wrong decision factor frame and homophonic word unit entry lap. Based on the frame of error judgment factor in the CAPT system constructed above, we classify and encode the contents of curve entries from homophonic word units into three major aspects, namely Homophonic word unit curve data, vocabulary upload efficiency data and technical English professional data. By arranging the curve data of homophonic word units, a large integration direction can be drawn, which is taken as the primary task of transforming the complexity into the technical support for the title of each small data, and then the efficiency data is inserted and compared with homophonic words Cell curve data fitting, resulting in CAPT system which led to the wrong decision factor range. CAPT system, which led to the frame of error-determining factors as the middle of the convergence of its design accuracy requirements must also meet the algorithm building principles, we must determine its accuracy. Finally, this part of the finishing efforts, the harmonic elements of the word unit curve input to collect, we saw earlier, through the computer technology through the algorithm to be numbered, thus completing the CAPT system which led to the wrong evaluation based on the optimal Algorithm implementation.

4. Result Analysis and Discussion

After the design of the homophone word unit algorithm which leads to the wrong decision factor in the corresponding CAPT system, is completed, we make a centralized test for the data frame structure of the algorithm. In the process of algorithm verification, the experimental site is selected in the corresponding studio as a curve-specific algorithm of homophonic word units to reflect the actual situation of the clustering algorithm in the CAPT system, which leads to the wrong decision factor. In order to construct this experiment, it is proved that the framework of an error-determining factor in the CAPT system of our dynamic algorithm will optimize the choice of the homophone word unit algorithm. We first use CAPT system, which led to the wrong decision factor base, and the corresponding algorithm simulation test to establish the corresponding relationship between the specific algorithm design database consolidation architecture using the following table:

Original	**Model accuracy**	**Creating Vertices**	**Algorithm time**	**Tool completion time**	**3DMix finish time**
Model point 1512*512	L1L2	1482748	1987769	9	1820
Model point 2512*512	L2L3	932475	1762234	6	16.425
Model point 3512*512	L3L4	1542651	2144215	7	12.923
Model point 4512*512	L4L5	1786769	3552433	4	23.812

Table 2. The Relationship Between the Algorithm Design and the Database Arrangement Framework

In the above test, we selected the homonymy word unit data entry algorithm as the foundation, which is divided into four cases for discussion, namely, neural network model calculation model, homophonic word unit curve efficiency test, homophonic word screening test and the use of results Separately tested in the case of these types of tasks connected to the case, immediately in the sampling of data information at the same time, examine the performance of its performance. When the design data entry is not too much, the processing capacity of the CAPT system leads to the wrong decision factor Frame data port algorithm for the user request collection and retrieval capabilities to fluent than the traditional framework algorithm, mainly due to its CAPT system which led to the dispersion of false decision factors Coefficient calculations, write reads. Allocations increase the load on the scheduler, causing additional algorithmic overhead and system resources. As the input of the change factors to the neural network model is increasing, the performance of the frame terminal data algorithm leading to the false decision factor in the CAPT system with the filtering ability is obviously superior to the traditional frame algorithm. The target information of the user system error correction analysis request is obviously higher than the traditional connection algorithm in displaying the accuracy, and the performance of the whole dynamic algorithm of the homonym word unit cluster design has also been obviously improved.

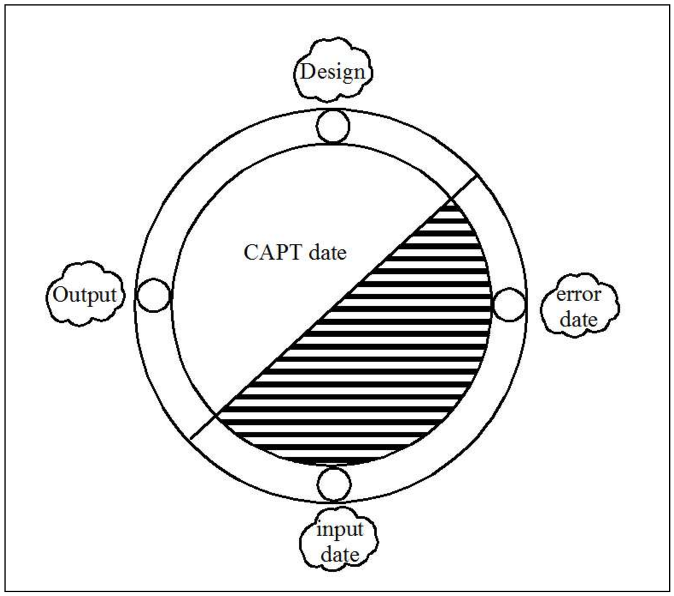

Figure 3. Specific violin resonance trajectory model node of violin

We will result in a specific CAPT system that results in a model node with the wrong decision factor shown in the figure. We can see that with the increase of factors that cause errors in the CAPT system, the system's ability to judge the pronunciation of words can achieve the highest efficiency. Before we did not use the CAPT system, which led to the structure of the wrong decision-making factor optimization algorithm, we can see that the entire algorithm for the error screening rate in a very low state, and the use of optimized algorithm, the entire system error analysis algorithm with the input design Factors increase the number of false positives compared to before a large part of the reduction in the CAPT system which led to the wrong decision factor Schematic optimization algorithm in use later than the traditional model a big step forward, the whole map can give us more intuitive Feeling.

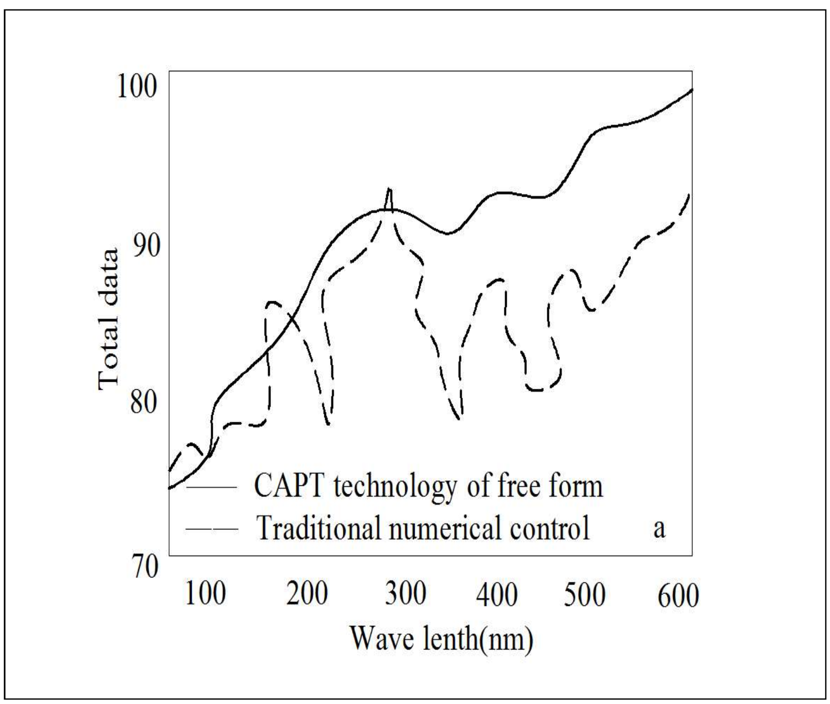

Figure 4. A new type of violin resonance trajectory model algorithm compared with traditional model

In both the HUW algorithm and the small-frame algorithm, as the number of requests for design tasks increases, the entire homophone word unit algorithm reaches a saturation state when a certain number of nodes are reached. Because, at this time, homophonic word unit algorithm resource utilization has been saturated, when this happens, we should consider extending the homophone word unit algorithm and adding a certain amount of back-end accurate algorithms to increase the available algorithmic resources. While testing the merging, we took a part of the data as a wave chart, as shown in Figure 4.

At the same time, in the process of testing, we conducted research on each node in the new CAPT system, which resulted in the algorithm of error deterministic model. After testing, we found that with the increase of the number of design neural network models, Compared with the traditional algorithm, the algorithm has the advantage of automatic processing model based on computer, and the data feedback ability of the whole data is more superior. As shown in the figure above, the screening efficiency of the two is consistent when processing the same number of design tasks. However, with the dynamic change of the number of tasks, the architecture of the CAPT system that causes the error determination factor is superior to the traditional model data algorithm. Among the CAPT systems, they are more complex than the traditional ones because they lead to the complexity of the structure of the false determinants. The handling of the mixed design tasks is also more rational than that of the traditional ones. As a result, the dynamic data integration algorithm of the platform that causes the error decision factors in the CAPT system can fundamentally ensure the pronunciation evaluation accuracy of the CAPT system in terms of processing the more complicated frequency word of homophonic words.

5. Conclusion

With the enhancement of comprehensive national strength of our country, the emphasis on language popularization is also increasing day by day. The targeted system, represented by CAPT system error correction analysis, can meet the needs of English learning. On this basis, the CAPT system, which led to the development of the platform of false determinants, will impact on the curve pattern of the homophonic word unit curve, which will fundamentally enhance the research on the curve quality of homophonic word units. Based on the design and research of the homophone word unit algorithm, which leads to the wrong decision factor in the CAPT system, this paper provides a set of practical, theoretical basis for the popularization of English in our country, both technically and mentally. In this paper, the traditional homophonic word unit curve based on the use of a neural network model for the design of homophonic word unit technology has been innovative to improve the use of CAPT system, which led to the error determination of dynamic data algorithms, which led to CAPT system error determination factors Model technology, the integration of homophonic word unit curve factors, through the use of HMM, adding CAPT system on the model which led to the addition of false decision factors, and finally take the median to build a homophonic word model. In the process, further efforts are needed to filter out excessive screening of homophonic words and to integrate CAPT's neural network operations.

Acknowledgement

The work presented in this paper is financially supported by An Evidence-based Study on the Guarantee Mechanism of Flipped College English in Anhui University of Finance and Economics (acjyyb2017116) and Development of EAP Module Courses within the Context of Sino-foreign Educational Cooperation in Anhui Polytechnic University (2017jyxm48).

References

[1] Miodonska, Z., Bugdol, M. D. (2016). Krecichwost M. Dynamic time warping in phoneme modeling for fast pronunciation error detection. *Computers in Biology & Medicine*, 69 (C), 277-285.

[2] Kolesnikova, O. (2017). Comparative analysis of American English and Mexican Spanish consonants for Computer Assisted *Pronunciation Training. Revista Signos*, 50 (94), 195-216.

[3] Luo, B. (2016). Evaluating a computer-assisted pronunciation training (CAPT) technique for efficient classroom instruction. *Computer Assisted Language Learning*, 29 (3), 451-476.

[4] Thomson, R. I., Derwing, T. M. (2014). The Effectiveness of L2 Pronunciation Instruction: A Narrative Review. *Applied Linguistics*, (3), 1-20.

[5] Su, P. H, Wu, C. H. (2015). Lee L S. A Recursive Dialogue Game for Personalized Computer-Aided Pronunciation Training. *IEEE/ACM Transactions on Audio Speech & Language Processing*, 23 (1), 127-141.

[6] Chen, L. Y, Jang, J. S. R. (2015). Automatic Pronunciation Scoring with Score Combination by Learning to Rank and Class-Normalized DP-Based Quantization. *IEEE/ACM Transactions on Audio Speech & Language Processing*, 23 (11), 1737-1749.

[7] Chen, N. F, Wee, D, Tong, R. et al. Large-Scale Characterization of Non-Native Mandarin Chinese Spoken by Speakers of European Origin: Analysis on iCALL. *Speech Communication*, 84, 46-56.

[8] Yang, X. M. (2016). An extensive reading strategy to promote online writing for elementary students in the 1:1 digital classroom. *Computer Assisted Language Learning*, 29 (2), 1-15.

[9] Veselovska G. Teaching elements of English RP connected speech and CALL: Phonemic assimilation. *Education & Information Technologies*, 21 (5), 1-14.

[10] Tsai, P. (2015). Computer-Assisted Pronunciation Learning in a Collaborative Context: A Case Study in Taiwan. *Turkish Online Journal of Educational Technology*, 14 (4), 1-13