A Pedagogical Experiment for Evaluation of Online English Courses Using the Principal Component Analysis (PCA)

Tatiana A. Kokodey
Sevastopol State University
Sevastopol, Russia, 299007
tanya.kokodey@gmail.com

Olga N. Golovko
Sevastopol State University
Sevastopol, Russia, 299007
pedagogical_education@mail.ru

Vitalina V. Khituschenko
Sevastopol State University
Sevastopol, Russia, 299007
vitalinkatheone@gmail.com

Vyacheslav A. Ley
Sevastopol State University
Sevastopol, Russia, 299007
peot2@mail.ru

Abstract

The focus of this research is to describe and analyze the results of a pedagogical experiment which sought to introduce online English-language courses into the educational process of Sevastopol State University (SevSU). In doing so, we evaluated the efficacy of these online English-language courses using the Principal Component Analysis (PCA) in the GRETL Statistical Software Package. That is, we created two random samples of 30 third-year undergraduate students in each. The first sample, the control group, consisted of students who study English using traditional methods without the use of distance learning technologies. The second sample is an experimental group, which included students whose English-language professional training from the fourth semester is provided using the Moodle online course in English. Then we created a set of initial data in GRETL for the experimental group, which consisted of the final grades of students in English for five semesters. Since the Moodle course was introduced for the experimental group in the fourth semester, we used the method of Principal Components Analysis to create two integral evaluation indicators calculated as the first principal components: before the introduction of Moodle and after that. Then we analyzed the dynamics of the difference between these values for the worst student of the experimental group. After the introduction of Moodle, the worst student improved his integrated result to 7.57. From the results of the calculations obtained, it can be concluded that with other permanent factors the reason for the improvement of learning achievement in the experimental group is the

In: Jože Rugelj, Maria Lapina (eds.): Proceedings of SLET-2019 – International Scientic Conference Innovative Approaches to the Application of Digital Technologies in Education and Research, Stavropol – Dombay, Russia, 20-23 May 2019, published at http://ceur-ws.org

use of the e-learning environment.

1 Introduction

Significant influence on the formation of the theory of pedagogical experiment was made by such authors as C. Hicks[Hic67], E.V. Yakovlev [Yak10], I.A. Stepankin[Step17], V.A. Stoff [Sto78] ,G.I. Batischev[Bat90], etc. Summarizing their definitions of this concept, we can agree with the following interpretation: a scientific experiment of transforming the pedagogical process under precisely measured conditions[Sol02]. However, despite significant developments in theoretical statements, the contemporary pedagogical science and practice are lacking a widely recognized unambiguous understanding of methodological foundations of a pedagogical experiment, as well as ways to efficiently implement it[Kle15] [Pol00].

2 Task

In this regard, the purpose of this article is, first, to test the hypothesis of the experiment conducted to estimate efficacy of online English courses, second, to expand the set of methods used to analyze pedagogical experiment data.

3 Development Of Methodology

3.1 Experiment Purpose

The purpose of the experiment is to assess the impact of the digitalization of the educational process at Sevastopol State University. In particular, of the introduction of e-learning English courses[Gal04], on the effectiveness of the English-language professional education of students at the bachelor's level, thus creating conditions for the further development of the e-learning environment at the university[Kor02] [Schr08].

3.2 Tasks And Location Of The Experiment

– Create two random samples of 30 third-year undergraduate students in each, so that different areas of study and different Institutes of the University are represented. The first sample, the control group, would consist of students who study English using traditional methods without the use of distance learning technologies. The second sample is an experimental group, which would include students whose English-language professional training from the fourth semester is provided using the Moodle online course in English.
– To generate a set of initial statistical data for both the control and experimental groups derived from recorded final grades in English for each student for each of the five semesters of study.
– Process the basic data in the GRETL statistical package using the Primary Component Analysis (PCA) tools to determine if the English language teaching inLMS Moodle is providing students with better performance in English.
– Analyze the results and make conclusions. The location of the experiment is Sevastopol State University.

3.3 Experiment Sample

Bachelor's degree students studying English are randomly selected from a variety of backgrounds. All students at the time of the experiment (January 10th, 2019) were studying in the third year (beginning of the sixth semester). Their total number is 60 people, 30 of whom are members of the control group, which used traditional methods of learning English (did not work in the system of Moodle), and 30 in the experimental group, who began using online English courses in Moodle at beginning of the fourth semester.

3.4 Hypothesis

The use of online English courses in LMS Moodlefor all full-time students, regardless of their specialty, increases their achievement in this discipline.

3.5 The Method Of Calculating The Integral Grade Indicator Using Principal Component Analysis (PCA)

The integral grade indicator can be obtained using Principal Component Analysis (PCA) [Kou07]in the Open Source Software – GRETL – by calculating the first principal component based on the aggregate valueofeach student's grades in time[Kal5]. That is, the PCA is applied in order to provide a generalization of initial grade metrics of students before and after online English courses were introduced in LMS Moodle[Dud08]. The principal component analysis, developed in 1901, is usually applied to compress excessive volumes of information for its easier interpretation[Rak99] . As far as initial indicators $x_1, \cdots, x_p$ are correlated with each other, it is possible to define new aggregated variables $y_1, \cdots, y_{p'}$ (y_j–principal component), $p' < p$. The new indicators $y_1, \cdots, y_{p'}$ are linear combinations of initial indicators $x_1, \cdots, x_p$, formula (1).

$$y_1(x) = w_{11}(\frac{x_1 - \overline{x_1}}{\sigma_1}) + \cdots + w_{51}(\frac{x_5 - \overline{x_5}}{\sigma_5}); \quad (1)$$

where $\overline{x_j}$ and σ_j – the average and standard deviation of x_j;
w_{j1} – coefficients of the most significant principal component ($\sum_{j=1}^{6} w_{j1}^2 = 1$);
y_1 – the most significant principal component that can be interpreted as the integral grade indicator. The value λ_1 is the maximum eigenvalue for the first principal component y_1.

4 Results

Below we will describe results of the experiment by stages. Stage 1.At the first stage, we created a set of initial data in GRETL software[Pol05] for the experimental group, which consists of the final grades of students in English for five semesters, Fig. 1. and 2. "Student" is a variable indicating the student's number, and the variables Semester1...Semester5 indicate the final grade for the corresponding semesters according to the 100-point system.

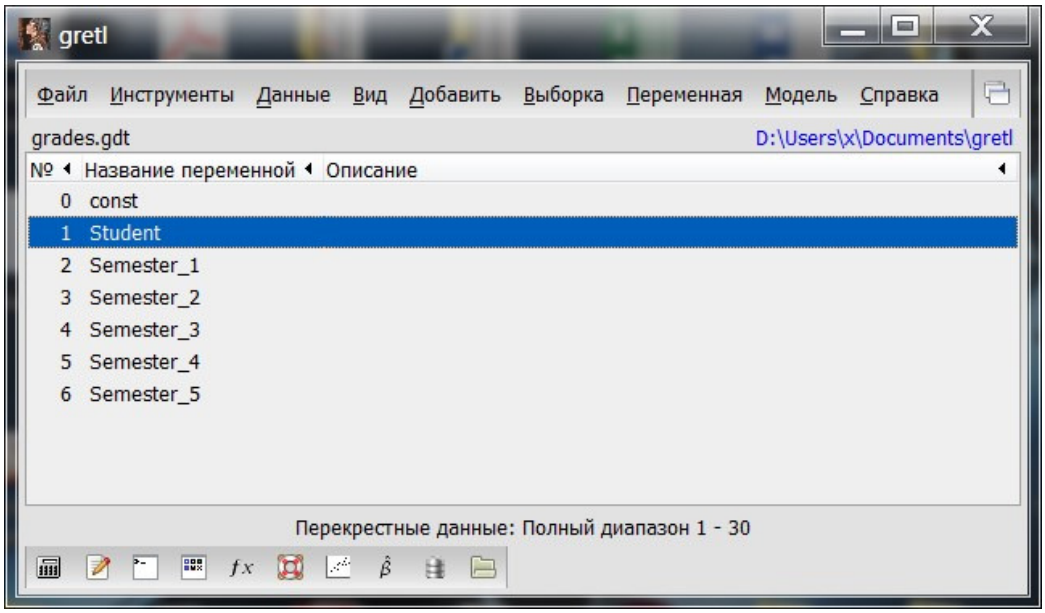
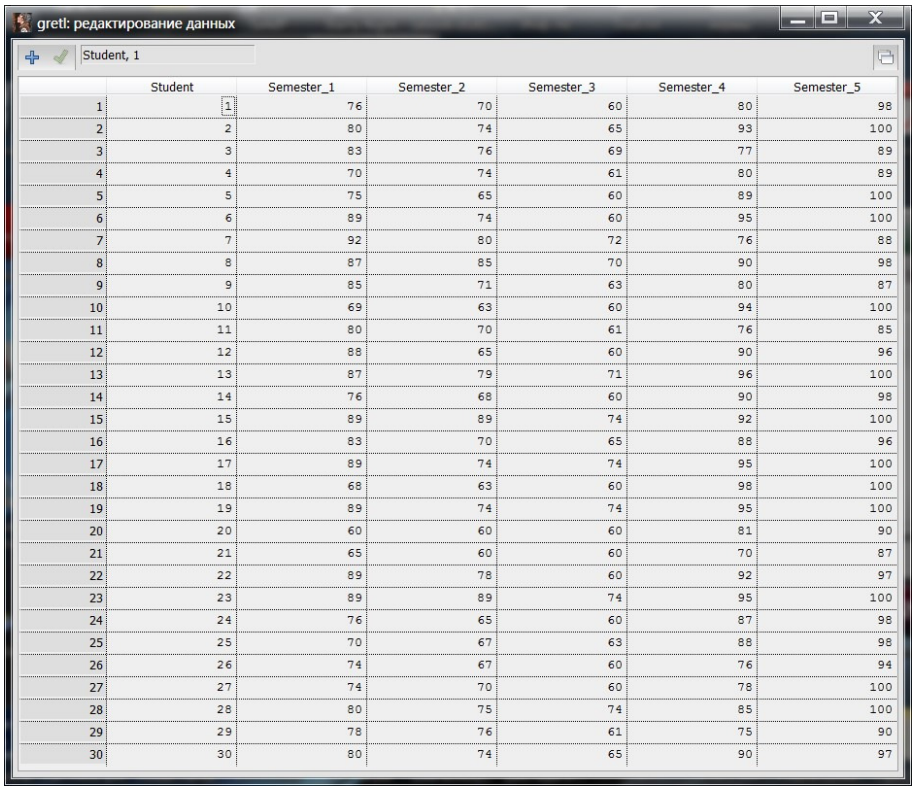


Figure 1: The initial data set for the experimental group in GRETL software

Stage 2: Since the Moodle course was introduced for the experimental group in the fourth semester, we use the method of principal components to create two integral evaluation indicators[Dyak09] [Mukh04] (calculated as the first principal components): before the introduction of Moodle (Y1_1) and after that (Y2_1).
Y1_1 builds on the space of the initial indicators: Semester_1, Semester_2 and Semester_3, taking into account the estimates for the first three semesters of English language teaching. The implementation of this stage in the GRETL software environment is shown in Fig. 3 On Fig.3,the first principal component Y1_1 has the designation



	Student	Semester_1	Semester_2	Semester_3	Semester_4	Semester_5
1	1	76	70	60	80	98
2	2	80	74	65	93	100
3	3	83	76	69	77	89
4	4	70	74	61	80	89
5	5	75	65	60	89	100
6	6	89	74	60	95	100
7	7	92	80	72	76	88
8	8	87	85	70	90	98
9	9	85	71	63	80	87
10	10	69	63	60	94	100
11	11	80	70	61	76	85
12	12	88	65	60	90	96
13	13	87	79	71	96	100
14	14	76	68	60	90	98
15	15	89	89	74	92	100
16	16	83	70	65	88	96
17	17	89	74	74	95	100
18	18	68	63	60	98	100
19	19	89	74	74	95	100
20	20	60	60	60	81	90
21	21	65	60	60	70	87
22	22	89	78	60	92	97
23	23	89	89	74	95	100
24	24	76	65	60	87	98
25	25	70	67	63	88	98
26	26	74	67	60	76	94
27	27	74	70	60	78	100
28	28	80	75	74	85	100
29	29	78	76	61	75	90
30	30	80	74	65	90	97

Figure 2: Displaying the source data of the Gretl software

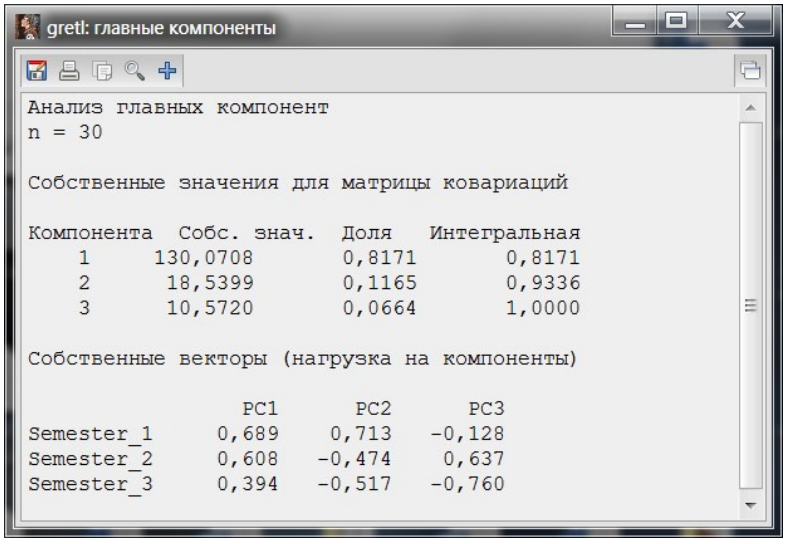


Figure 3: The results of modeling the integrated assessment indicator in English before the introduction of Moodle (first three semesters of training)

PC1, it has the highest significance and is considered as an integral (generalized) evaluation indicator, formula (2) :

$$Y_{1\_1} = 0,689 * Semester_1 + 0,608 * Semester_2 + 0,394 * Semester_3 \quad (2)$$

The formula (2) shows the dependency of the integral index Y1_1 on the initial indicators of Semester1...Semester3. Equation coefficients (2) show the contribution of each individual index to the inte-

gral indicator Y1..1.
The columns of the PC_i in the modeling results window (Fig. 3) contain values of the coefficients of the principal components $w_j = (w1_j, \cdots, w_pj)'$, and according to the column PC1 the first principal component Y was developed.
The contribution of PC_1 (Y1_1) to the total variance of individual indices of the estimations is maximal and in absolute terms is equal to $\lambda_1 = 130$, and in percentage - 81,71%. Therefore, the first principal component Y1_1 (PC1) can be considered the integral grade indicator. The other principal components of PC2 and PC3 with insignificant contributions to the overall variance can be ignored. For each of the thirty students in the experimental group, we will calculate the values of Y1_1 by formula (2). The result of the calculations is shown in Fig. 4. Integral index Y2_1 is built on the basis of the initial indicators Semester_4, Semester_5, taking into

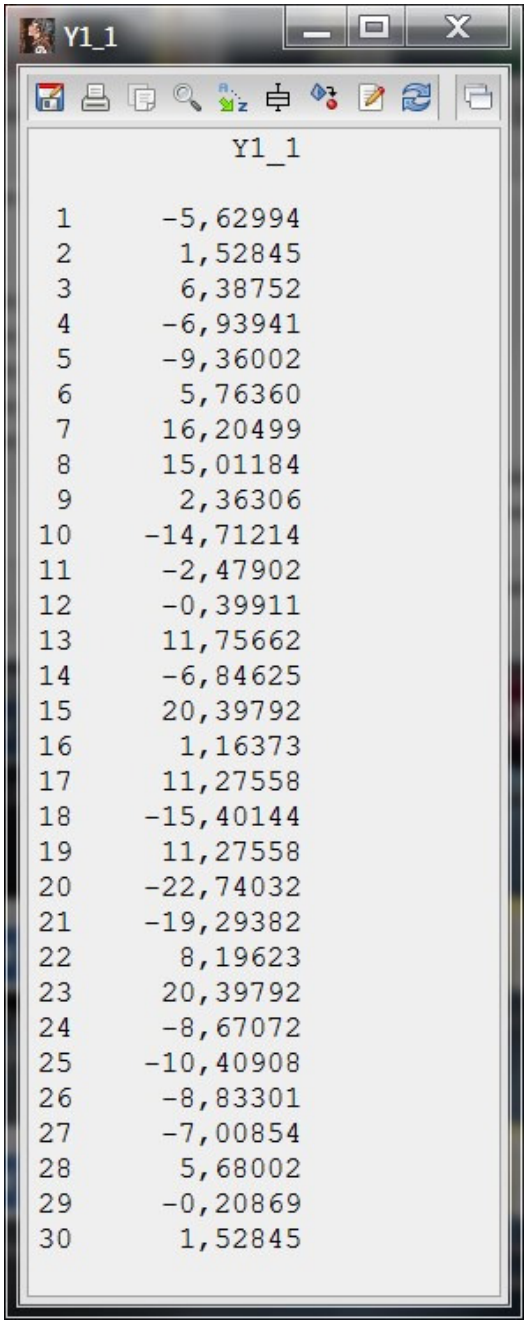


Figure 4: The value of the main component Y_{1_1}

account the estimates for the last two semesters of English language teaching in Moodle. The realization of this stage in the GRETL software environment is shown in Fig. 5. Based on the results of modeling of Fig.5, we will

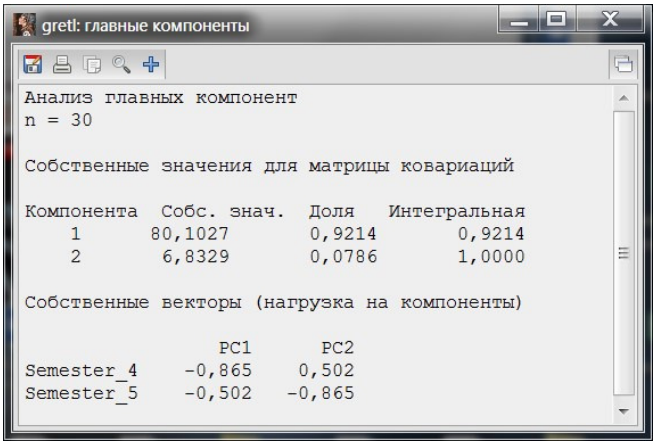


Figure 5: The results of modeling the integral assessment indicator for English language after the introduction of Moodle (last two semesters of training)

construct an integral index of Y2_1 evaluation as the first principal component, formula (3):

$$Y_{2\_1} = 0,865 * Semester_4 + 0,502 * Semester_5 \tag{3}$$

This is possible because the PC1 (Y2_1) contribution to the total variance is enough and amounts to 80.1%. The values of this principal component are shown in Fig.(6)

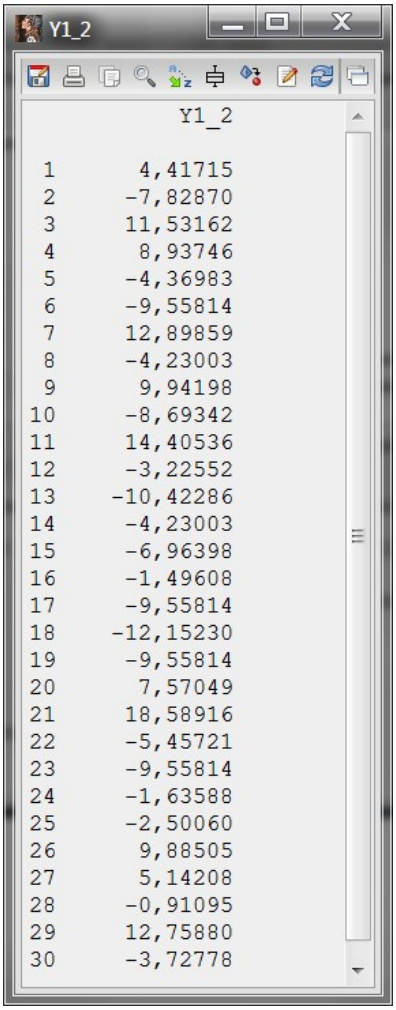


Figure 6: The value of the principal component Y2_1

5 Discussion

We analyzed the dynamics of the integral grade indicator in English language as a difference of values Y1_1 and Y2_1 for the worst student of the experimental group. In the first three semesters the best results were shown by students 15 and 23, showing the integral result Y1_1= 20.4 (Fig. (5). The worst result on this indicator (-22.74) was shown by student 20. After the introduction of Moodle, the worst student improved his integrated result (two semesters) to 7.57, which may be the result of using an online Englishcourse[Sal09][Kle15]. This statement is supported by the visual analysis of the experimental grouppresented by the box diagram in Fig. (7)

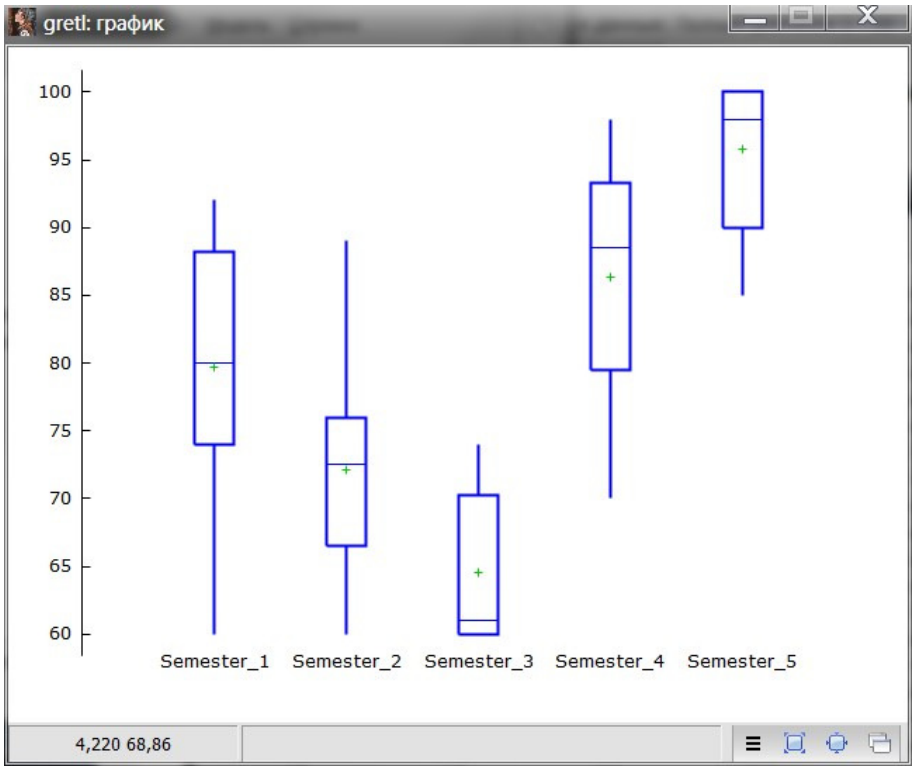


Figure 7: Dynamics of the meangrade for the experimental group

In the first three semesters, the means, as well as medians decrease, and after the introduction of e-learning tend to increase. In the control group, such dynamics were not indicated.

6 Conclusion

From the results of the calculations obtained, it can be concluded that with other permanent factors the reason for the improvement of learning achievement in the experimental group is the use of the e-learning environment. In other words, the use of online courses is a factor in improving English learning achievement. This research was supported by the grant 19-010-00377 of the RFBR (Russian Foundation for Basic Research), Developing a strategic management system for the digital education enhancement in the Russian Federation.

References

[Hic67] C. Hicks Basic principals of experiment planning. *M.,Mir*, 1967.

[Yak10] E. V. Yakovlev, N. O. Yakovleva Pedagogical research:content and presentation of results *Chelyabinsk.:RBIU*, 2010.

[Step17] I. A. Stepankin Features of the organization of pedagogical experiment. *Modern scientist*, 6 2010.

[Sto78] V. A. Stoff Problems of the methodology of scientific knowledge. *M.:Higher School*, 1978.

[Bat90] G. I. Batishchev Pedagogical experimentation. *Sov. Pedagogy*, 1990.

[Kou07] T. Koufel Econometrics. Problem solving with the use of GRETL. *Moscow: Hot line - Telecom*, 2007.

[Dud08] G. Dudeney, N. Hockly How to Teach English with Technology. *Pearson Longman*, 2008.

[Schr08] J. A. Schrader Information processes and information environment. *Scientific and technical information. Ser. 2, Information processes and systems*, 9:3–7, 2008.

[Sol02] V. I. Soldatkin Fundamentals of open education. *M.: NIIU RAO*, 2002.

[Sal09] L. K. Salnaya Training for the professionally oriented foreign-language communication. *Taganrog: TTI SFU*, 2009.

[Rak99] E. A. Rakitina. Information fields in educational activity. *Informatics and Education*, 1, 1999.

[Kle15] N. P. Kleinosova. Design and development of a distance learning course in the Moodle 2.7 environment: teaching materials. *Ryazan. state. radio engineering. Ryazan*, 2015.

[Pol00] E. S. Polat Project method for foreign language lessons. *Foreign languages at school*, 2–3, 2000.

[Pol05] E. S. Polat New pedagogical and information technologies in the education system: textbook *M.: Academy*, 2005.

[Mukh04] S. A. Mukhina Alternative pedagogical technologies in education. *Rostov-on-Don: Phoenix Publishing House*, 2004

[Kor02] N. F. Koryakovtseva. Modern methods of the organization of the independent work of the foreign language students. *MOSCOW; ARKTI*, 2002.

[Kal5] K. A. Kalyuzhny. Information environment and information environment of science:Essence and purpose. *Science. Innovations. Education*,18:7–23, 2015.

[Dyak09] V. P. Dyakonov New information technologies: textbook. *MOSCOW: SOLON-PRESS*, 2009.

[Dyak09] A. V. Goncharov, Y. A. Medvedev We are studying the distance learning system Moodle/ Part 1 *Vladimir: VlSU*, 2011.

[Gal04] N. D. Galskova Modern methods of teaching foreign languages *M: ARKTI*, 2004.