

CLASSES EVALUATION – METHODS AND TOOLS

Tadeusz Grabiński*, Andrzej Frycz-Modrzewski Krakow University
Barbara Barszcz-Przełożny, Andrzej Frycz-Modrzewski Krakow University
Zbigniew Paszek, Andrzej Frycz-Modrzewski Krakow University
Szczepan Urlik, Andrzej Frycz-Modrzewski Krakow University

Abstract: This study presents a method, tools, course and results of foreign language classes evaluation conducted in the summer semester 2012/2013 in the Andrzej Frycz – Modrzewski Krakow University. Because a new evaluation procedure has been implemented at the University, the former method – based on paper forms filled in by the students – was abandoned. On the surveyanyplace.com website, a free account has been registered and the form of the evaluation questionnaire has been inserted. This coverage presents results of a taxometric analysis aimed at checking the degree of mutual correspondence (correlation) between certain criteria and instancing a graphic presentation of the evaluation results in a multidimensional perspective. In order to classify the grading criteria, the Ward's agglomerative method, along with Euclidean metric as a measure of criteria similarity, have been used. Calculations have been made with the use of Statistica package. Results of the questionnaire show that foreign language teaching at the Andrzej Frycz Modrzewski Krakow University is conducted professionally and on a high factual level.

Key words: evaluation, quality management in education, a questionnaire form, taxometric analysis, correlation, average grade distribution matrix, factor correlation matrix, teacher's four – dimensional profiles.

JEL classification: C39, C52, I20

EVALUACIJA ČASOVA - METODE I SREDSTVA

Sažetak: Ova studija predstavlja metod, alatke, tok i rezultate evaluacije nastave stranog jezika sprovedene u letnjem semestru 2012/2013 na Krakovskoj akademiji „Andrej Frič Modževski“. Zbog toga što je novi postupak evaluacije sproveden na Univerzitetu, raniji metod - zasnovan na papirnim obrascima popunjenim od strane studenata – je napušten. Na sajtu surveyanyplace.com, registrovan je besplatan nalog i umetnut je obrazac upitnika za evaluaciju. Ovaj rad obuhvata rezultate taksometrijske analize u cilju provere stepena međusobne korespondencije (korelacije) između

* tg@uek.krakow.pl

određenih kriterijuma instanciranjem grafičkog prikaza rezultata evaluacije u multidimenzionalnoj perspektivi. Da bi se klasifikovali kriterijumi ocenjivanja, koristili su se Vardov aglomerativni metod sa Euklidovom metrikom kao merom sličnosti kriterijuma. Izračunavanje je izvršeno uz korišćenje paketa Statistica. Rezultati upitnika pokazuju da se nastava stranog jezika na Krakovskoj akademiji „Andrej Frič Modževski“ sprovodi profesionalno i na činjenično visokom nivou.

Ključne reči: *evaluacija, upravljanje kvalitetom u obrazovanju, obrazac upitnika, taksometrijska analiza, korelacija, matrica raspodele prosečne ocene, matrica korelacija među faktorima, četvorodimenzionalni profili nastavnika.*

1. INTRODUCTION

This study presents a method, tools, course and results of foreign language classes evaluation conducted in the summer semester 2012/2013 at the Andrzej Frycz – Modrzewski Krakow University. Because a new evaluation procedure has been implemented at the University, the former method – based on paper forms filled in by the students – was abandoned. It was essential to create a new evaluation system that would be an element of the education quality system.

The following variants of conducting the evaluation procedure have been considered:

1. Internet evaluation, based on making the evaluation form available on the virtual dean's office website and calling on students to take part in the evaluation.
2. Internet evaluation based on devoting a part of the last classes (15 – 25 min) to fill in the evaluation form available on the Internet.
3. Evaluation with the use of mobile devices, and to be more specific – text messages sent by the students during the last classes from their mobile phones; the messages would include answers to questions projected on a slide.

Undoubtedly, the best solution is the first of the above mentioned procedures (Patton, 1987). Unfortunately, because of formal and legal reasons, it is impossible to make students participate in a questionnaire, e.g. conditioning passing a semester on taking part in the evaluation. As a result, the so – called 'reflexivity' of the questionnaire is quite small in this case, which, when there are few answers, makes their analysis pointless, because of the lack of the data's representative nature (Millman, & Darling-Hammond, 1990).

The next two solutions guarantee reflexivity of opinions on the attendance level. Because these are the last classes of the semester, when teachers often give final grades, attendance on those classes is relatively high.

Each of the two procedures has minor shortcomings. In the second variant, Internet access is essential. If classes take place in a classroom without

computers, it is necessary to book a lab, an Internet café or a room in the library.

In the third variant, during the last but one class students must be warned about the evaluation and asked to bring their mobile phones with them. It may be argued that additional costs (text messages) are imposed on students. This idea may be considered arguable by pointing out that these are very small costs, much lower than e.g. the costs connected with travelling to school, buying books, materials for notes, etc., which students must take without expecting any compensation.

There are also possible solutions which are a combination of the above mentioned variants. For example, joining the Internet procedure (variant 1) with one of the two remaining variants. The drawbacks of those procedures are difficulties connected with identifying those students who would like to hand back two opinions on one teacher.

In the discussed case, the second variant of evaluation procedure was chosen and treated as an experiment. Next semester, it is planned to conduct the evaluation using the third procedure. Only a comparative analysis of the two approaches will allow to judge which of these variants is going to be chosen as the standard one.

This coverage presents results of a taxometric analysis aimed at checking the degree of mutual correspondence (correlation) between certain criteria and instancing a graphic presentation of the evaluation results in a multidimensional perspective. In order to classify the grading criteria, the Ward's agglomerative method, with Euclidean metric as a measure of criteria similarity, have been used. Calculations have been made with the use of Statistica package.

The survey in 2013 lasted six weeks, starting from the middle of May. Altogether, 1390 students' opinions on 15 teachers have been gathered (more than 90 questionnaire forms for one teacher, on average). Range of the number of opinions characterizing specific teachers is considerable and is included in the bracket between 244 and 37. It must be mentioned that about 10% of the opinions have been eliminated; those were opinions for which the questionnaire code did not fit the password, and also opinions for two teachers which were small in quantity (below 30).

2. EVALUATION TOOLS

As the opinion gathering tool, an application available on <https://surveyanyplace.com> was chosen, mainly because it is possible to use it on mobile devices. The website contains detailed instructions on creating forms

and monitoring the course of acquiring, saving and analyzing answers given by the surveyed.

The chosen tool has achieved its aim. The application did not cause any problems, neither on the stage of creating the electronic form, nor while conducting the evaluation (logging – in, saving and exporting data).

The second element of the evaluation was the questionnaire form. It consists of three parts:

- 10 questions evaluating teacher's work, each with an identical 5-point grading scale (during the analysis letter answers A, B, C, D, E are changed into numbers 5, 4, 3, 2, 1),
- 2 questions concerning usefulness of the e-workbook (if it was used during classes),
- a part containing general descriptive comments.

Apart from the answers to the above mentioned 12 questions, students inserted into the form a code identifying the teacher, the student group (year, the field of studies, form and degree of studies). The code has a form of a 6-element sequence of small letters, e.g. 'abcdef', and is a synthetic identifier of a class. In other words, it is an equivalent of certificate questions. The advantage of using codes instead of certificate questions is the possibility of exercising easier control over the data correctness. **The system allows to fill in the form only after congruity of the code inserted by the student with the access code provided by the person who conducts the evaluation has been verified. The password is removed by the person conducting the evaluation after all the students have completed the evaluation process.**

Subsequent symbols of the code have the following meaning:

a_1 – language , e.g. a=angielski (English), n=niemiecki (German), etc.

a_{23} – teacher's initials, e.g. js=John Smith

a_4 – field of studies, e.g. a=administration, b=architecture

a_5 – form of studies (full – time, part – time), enrolment year (11/12, 12/13..) and semester (winter / summer), e.g. a=full – time studies, 11/12 year group, winter semester

a_6 – degree of studies (undergraduate, graduate, uniform Master's degree studies) and the student group number, e.g. a=undergraduate level, group 1

As a result, the data record contained 22 pieces of information detailed in the table. The first part contains identification information used as classification criteria in result tables:

- date and time of completing the questionnaire,
- name of the teacher,
- language (English, German),

- field of studies,
- form of studies (full – time, part – time),
- degree of studies (undergraduate, graduate).

In the second part of the data record, there are answers to individual questions covered by the form, as well as an arithmetic mean of the grades given in the first nine questions of the evaluation questionnaire. When calculating the mean, two last questions connected with only a part of classes, and question 10 which concerns evaluating the teachers' work in a smaller degree – were not taken into consideration.

The questionnaire form was saved as an electronic document, according to the demands of the Internet application. Graph 1 shows, in the form of screenshots, the welcome page of the e-form as well as a few example questions from the questionnaire.


The person filling in the form may change their answers until they click the FORM COMPLETE icon, situated on the last screen of the e- form.

Table 1

Record structure with source data

Nok	Specification	Inf. type
K1	Survey date	h-m-s y-m-d
K2	Group code	Text
K3	Teacher	Text
K4	Language	Letter
K5	Field of studies	Letter
K6	Full time / part time	Letter
K7	Complementary master's studies / undergraduate	Letter
K8	Mean of 9 questions (3-11)	Number
1	Teacher's punctuality	A=5...E=1
2	Pace and dynamics of the classes	A=5...E=1
3	Clarity of teacher's commands	A=5...E=1
4	Teacher's commitment and readiness to provide help or explanation	A=5...E=1
5	Method of conducting classes	A=5...E=1
6	Motivating students to learn foreign languages	A=5...E=1
7	Atmosphere of the classes	A=5...E=1
8	Fairness in assessment	A=5...E=1
9	Effectiveness of the classes	A=5...E=1
10	Keeping balance between gram mar and practical language use	A=5...E=1
11	Are you a student who uses Total English online workbook	T=Yes N=No
12	How does using the platform influence language learning attractiveness	A=5...E=1
13	How does the possibility of self – study influence the effectiveness of language learning	A=5...E=1
14	Comments, suggestions and proposals that would help improve foreign language teaching at AFM Krakow University	Text

Note: Compiled by author.



Krakowska Akademia
 im. Andrzeja Frycza Modrzewskiego

Witam

Kwestionariusz oceny lektoratu Niniejsza ankieta na służyć usprawnieniu i udoskonaleniu nauczania języków obcych w SJO KA

Imię (opcjonalnie)

e-mail (opcjonalnie)

Start >

Pytanie 2: Podaj kod lektora

< Wstecz **Dalej >**

Pytanie 3: Punktualność lektora

☐ A
☐ B
☐ C
☐ D
☐ E

A- ocena najwyższa, E-ocena najniższa

< Wstecz **Dalej >**

Pytanie 14: Dot. wyłącznie studentów korzystających z elektronicznego workbooka Total English online. W jakim stopniu możliwość samodzielnej nauki przy użyciu platformy wpływa na efektywność nauczania języka

☒ A
☐ B
☐ C
☐ D
☐ E

A- ocena najwyższa, E-ocena najniższa

< Wstecz **Dalej >**

Pytanie 15: Uwagi, sugestie i propozycje, umożliwiające usprawnienie nauczania języków obcych w KAAFM.

< Wstecz **Ankieta kompletna >**

Figure 1. Screenshots of chosen pages of the questionnaire form.

3. THE FORM OF PRESENTING EVALUATION RESULTS

Evaluation results are presented in four forms:

1. Individual opinions.
2. Total opinions (synthetic).
3. Opinions in disaggregated cross - sections (fields of study, years, forms of study, etc.)
4. Advanced forms of opinion presentation.

3.1. INDIVIDUAL OPINIONS

Detailed evaluation results for each question have been presented in tables containing layouts of students' answers for every teacher, a total of given answers [N] and an arithmetic mean of the grades given. When pointing the means, a 'school' scale was used, i.e. the highest grade [A] is assigned [5] points, the 'good' grade [B] = [4] points, etc. The equivalent of the lowest [E] is [1].

Teachers in tables are arranged according to declining values of average grades. Moreover, in a separate line in the tables, total averages (set on the basis of all 1390 students' opinions) are quoted. That is why it is easy to see which teacher is below or above the average level of students' grades.

Table 2 shows example collation of evaluation results for question 1 concerning punctuality in conducting lessons. If it is necessary, data from the tables may be presented in the form of graphs.

3.2. OPINIONS IN DISAGGREGATED CROSS – SECTIONS

Pieces of information included in the questionnaire code allow to set parameters that characterize opinions in more disaggregated student groups sectioned according to:

- form of studies (full –time, part- time),
- degree (undergraduate and graduate),
- field of studies, student group, language being taught, etc.

For example, in tables 4 and 5 information for the first both variants was given. If a teacher had classes with students from only one form or one degree of studies, his/her name appears only once. Therefore, the number of lines in the tables varies. To make interpretation easier, parameters for those teachers who have classes with full time and first degree studies were marked yellow. The last line of the tables presents differences between average grades of the teachers who teach full – time and part – time students, as well as between those who have classes with students of first and second degree studies.

The quoted tables show that from the point of view of all the criteria, full – time students are more demanding than part – time students, and graduate students and more demanding than undergraduate students.

Table 2

Example table with evaluation results for question concerning punctuality

Teachers/grades	5=A	4=B	3=C	2=D	1=E	N	Average
Teacher A	76	1				77	5,0
Teacher B	91		1			92	5,0
Teacher C	241	2			1	244	5,0
Teacher D	47	2				49	5,0
Teacher E	118	4		1		123	4,9
Teacher F	81	2	3			86	4,9
Teacher G	33	4				37	4,9
Teacher H	43		1		1	45	4,9
Teacher I	71	3	4			78	4,9
Total	1208	115	40	16	11	1390	4,8
Teacher J	49	5	3	1		58	4,8
Teacher K	114	12		2	4	132	4,7
Teacher L	24	12	1			37	4,6
Teacher M	103	36	10	5	2	156	4,5
Teacher N	42	18	5	2	1	68	4,4
Teacher O	75	14	12	5	2	108	4,4

Note: Compiled by author.

4. TAXONOMIC ANALYSIS OF THE EVALUATION RESULTS

The data from the evaluation questionnaires can be analyzed taxonomically, which will check the extent of the correlation of the evaluation criteria and the graphic presentation of the evaluation results with a multidimensional approach.

The starting point of the analysis is the matrix of average ratings, as determined from the particular criteria (questions) in the questionnaire (table 6). In the following tables and illustrations symbols from X1 to X13 have been introduced to mark the subsequent evaluation criteria, as well as the symbol X to mark the average rating of the first 9 criteria. Moreover, the number of opinions referring to particular teachers – N – has been adopted as an additional feature in the taxonomic analysis.

On the basis of the data from table 6 two elements have been determined: a matrix of correlation factors among the particular evaluation criteria (table 7) and a matrix of probability that a given correlation factor is insignificantly different from zero (table 8). The correlation factors significantly different from

zero, showing associated features (positively or negatively), have been marked in red in table 7. In table 8 zero values correspond with these factors.

Table 3

Average grades according to partial criteria and aggregated measures (general average, min – max values, spread)

Teacher	Commitment	Punctuality	Fairness	Atmosphere	Clarity	Average	Pace	Conduct	Effectiveness	Motivating	No of answers	Max	Min	Spread
Teacher A	5	5	5	5	5	4,9	4,9	5	4,9	4,9	244	5	4,9	0,1
Teacher B	4,9	5	4,9	4,9	4,9	4,9	4,9	4,9	4,9	4,9	92	5	4,9	0,1
Teacher C	5	5	5	5	4,9	4,9	4,8	4,9	4,9	4,9	77	5	4,8	0,2
Teacher D	5	4,9	4,8	4,9	4,9	4,9	4,9	4,8	4,8	4,8	123	5	4,8	0,1
Teacher E	4,8	4,9	4,8	4,8	4,8	4,8	4,8	4,7	4,7	4,6	86	4,9	4,6	0,3
Teacher F	4,8	4,9	4,6	4,8	4,6	4,7	4,8	4,7	4,8	4,6	37	4,9	4,6	0,3
Teacher G	4,9	5	4,7	4,8	4,8	4,7	4,7	4,5	4,4	4,2	49	5	4,2	0,7
Teacher H	4,7	4,4	4,8	4,9	4,7	4,7	4,5	4,6	4,6	4,6	108	4,9	4,4	0,4
Teacher I	4,8	4,8	4,7	4,8	4,6	4,6	4,6	4,7	4,5	4,4	58	4,8	4,4	0,4
Teacher J	4,8	4,4	4,7	4,6	4,7	4,6	4,5	4,5	4,5	4,7	68	4,8	4,4	0,3
Teacher K	4,7	4,9	4,7	4,8	4,6	4,6	4,4	4,4	4,4	4,3	78	4,9	4,3	0,5
Teacher L	4,8	4,6	4,8	4,5	4,6	4,5	4,3	4,2	4,3	4,4	37	4,8	4,2	0,6
Teacher M	4,7	4,5	4,7	4,8	4,6	4,5	4,2	4,3	4,3	4,2	156	4,8	4,2	0,5
Teacher N	4,8	4,9	4,3	4,5	4,3	4,5	4,5	4,4	4,2	4,2	45	4,9	4,2	0,7
Teacher O	4,5	4,7	4,5	4,1	4,3	4,2	4,1	3,8	3,9	3,8	132	4,7	3,8	0,9
In all	4,8	4,8	4,8	4,8	4,7	4,7	4,6	4,6	4,6	4,5	1390	4,8	4,5	0,3
Max	5	5	5	5	5	4,9	4,9	5	4,9	4,9		5	4,9	0,1
Min	4,5	4,4	4,3	4,1	4,3	4,2	4,1	3,8	3,9	3,8		4,5	3,8	0,7
Spread	0,4	0,6	0,7	0,9	0,6	0,7	0,9	1,1	1	1,1		1,1	0,4	0,7

Note: Compiled by author.

Table 4

Average grades according to partial criteria and aggregated measures (general average, min – max values, spread) set on the basis of full –time students (yellow colour) and part – time students

Teacher	Commitment	Punctuality	Fairness	Atmosphere	Clarity	Average	Pace	Conduct	Effectiveness	Motivating	No of answers	Max	Min	Spread
Teacher A	5	5	5	5	5	4,9	4,9	5	4,9	4,9	244	5	4,9	0,1
Teacher B	4,9	5	4,9	4,9	4,9	4,9	4,9	4,9	4,9	4,9	92	5	4,9	0,1
Teacher C	5	4,9	5	5	5	4,9	4,9	4,9	5	4,7	9	5	4,7	0,3
Teacher D	5	5	5	5	4,9	4,9	4,8	4,9	4,9	4,9	77	5	4,8	0,2
Teacher E	4,9	4,9	4,9	4,9	4,9	4,9	4,9	4,8	4,8	4,8	113	4,9	4,8	0,1
Teacher F	5	5	4,5	5	4,7	4,9	5	4,9	4,7	4,9	10	5	4,5	0,5
Teacher G	5	5	4,8	4,8	4,8	4,8	4,9	4,8	4,6	4,8	18	5	4,6	0,4
Teacher H	4,8	4,9	4,8	4,8	4,8	4,8	4,8	4,7	4,7	4,6	86	4,9	4,6	0,3
Teacher I	4,9	4,9	4,8	4,7	4,7	4,7	4,6	4,7	4,6	4,6	14	4,9	4,6	0,4
Teacher J	4,8	4,9	4,6	4,8	4,6	4,7	4,8	4,7	4,8	4,6	37	4,9	4,6	0,3
Teacher K	4,7	4,4	5	5	4,6	4,7	4,6	4,6	4,5	4,7	25	5	4,4	0,6
Teacher L	4,7	4,4	4,8	4,8	4,7	4,6	4,5	4,7	4,6	4,6	83	4,8	4,4	0,4
Teacher M	4,8	4,7	4,7	4,8	4,6	4,6	4,6	4,7	4,4	4,3	44	4,8	4,3	0,5
Teacher F	4,8	5	4,7	4,7	4,8	4,6	4,7	4,5	4,3	4,2	40	5	4,2	0,8
Teacher G	4,8	4,4	4,7	4,6	4,7	4,6	4,5	4,5	4,5	4,7	68	4,8	4,4	0,3
Teacher H	4,7	4,9	4,7	4,8	4,6	4,6	4,4	4,4	4,4	4,3	78	4,9	4,3	0,5
Teacher I	4,7	4,4	4,7	4,9	4,8	4,6	4,3	4,4	4,4	4,4	55	4,9	4,3	0,6
Teacher J	4,7	4,6	4,8	4,6	4,6	4,5	4,4	4,3	4,5	4,4	30	4,8	4,3	0,5
Teacher K	4,6	4,5	4,7	4,7	4,5	4,4	4,1	4,3	4,2	4,1	101	4,7	4,1	0,6
Teacher L	5	4,6	4,7	4,4	4,7	4,3	3,9	3,6	3,6	4	7	5	3,6	1,4
Teacher M	4,6	4,7	4,5	4,1	4,4	4,3	4,1	3,9	4	3,9	37	4,7	3,9	0,8
Teacher N	4,6	4,8	4	4,3	4	4,2	4,3	4,1	3,9	3,7	27	4,8	3,7	1
Teacher O	4,5	4,8	4,5	4,1	4,3	4,2	4	3,8	3,9	3,8	95	4,8	3,8	1
Full time – Part time	-0,2	-0,2	-0,2	-0,3	-0,3	-0,3	-0,4	-0,5	-0,5	-0,5	-570	-0,2	-0,5	0,3

Note: Compiled by author.

Table 5

Average grades according to partial criteria and aggregated measures (general average, min – max values, spread) set on the basis of students of the I degree studies (yellow colour) and students of the II degree studies

Teacher	Commitment	Punctuality	Fairness	Atmosphere	Clarity	Average	Pace	Conduct	Effectiveness	Motivating	No of answers	Max	Min	Spread
Teacher A	5	5	5	5	5	5	5	5	5	4,9	196	5	4,9	0,1
Teacher B	5	5	4,7	5	5	4,9	5	5	5	4,8	11	5	4,7	0,3
Teacher C	5	4,8	4,9	5	5	4,9	4,9	5	5	4,9	23	5	4,8	0,2
Teacher D	4,9	5	5	4,9	4,9	4,9	4,9	4,9	4,9	4,9	81	5	4,9	0,1
Teacher E	5	5	5	5	4,9	4,9	4,8	4,9	4,9	4,9	77	5	4,8	0,2
Teacher F	5	4,8	5	5	5	4,9	4,8	4,7	4,9	4,9	10	5	4,7	0,3
Teacher G	5	4,9	4,8	4,9	4,9	4,9	4,9	4,8	4,8	4,8	123	5	4,8	0,1
Teacher H	4,9	4,9	4,8	4,8	4,9	4,8	4,7	4,8	4,8	4,8	48	4,9	4,7	0,2
Teacher I	4,9	5	4,8	4,9	4,8	4,8	4,8	4,8	4,7	4,7	32	5	4,7	0,3
Teacher J	4,8	4,9	4,8	4,9	4,8	4,8	4,8	4,7	4,7	4,7	62	4,9	4,7	0,2
Teacher K	5	5	4,7	4,8	4,8	4,8	4,7	4,8	4,8	4,5	6	5	4,5	0,5
Teacher L	4,8	4,9	4,6	4,8	4,6	4,7	4,8	4,7	4,8	4,6	37	4,9	4,6	0,3
Teacher M	4,9	4,8	4,8	4,6	4,8	4,7	4,6	4,7	4,7	4,4	24	4,9	4,4	0,5
Teacher F	4,9	5	4,7	4,8	4,8	4,7	4,7	4,5	4,4	4,2	49	5	4,2	0,7
Teacher G	4,8	4,7	4,8	4,8	4,6	4,6	4,6	4,7	4,4	4,4	52	4,8	4,4	0,4
Teacher H	4,8	4,4	4,7	4,6	4,7	4,6	4,5	4,5	4,5	4,7	68	4,8	4,4	0,3
Teacher I	4,7	4,3	4,8	4,8	4,6	4,6	4,4	4,5	4,5	4,6	85	4,8	4,3	0,5
Teacher J	4,8	4,6	4,8	4,5	4,6	4,5	4,3	4,2	4,3	4,4	37	4,8	4,2	0,6
Teacher K	4,6	4,5	4,7	4,7	4,6	4,5	4,2	4,3	4,3	4,2	146	4,7	4,2	0,6
Teacher L	4,8	4,9	4,3	4,5	4,3	4,5	4,5	4,4	4,2	4,2	45	4,9	4,2	0,7
Teacher M	4,6	4,8	4,7	4,7	4,4	4,4	4,1	4,2	4,2	4,1	46	4,8	4,1	0,7
Teacher N	4,7	4,9	4,6	4,3	4,5	4,3	4,2	4	4	3,9	100	4,9	3,9	1
Teacher O	4,2	4,3	4,1	3,5	4	3,8	3,7	3,3	3,5	3,6	32	4,3	3,3	1
Undergraduate - II degree studies	-0,1	-0,1	0	0	-0,1	-0,1	-0,1	-0,1	-0,1	-0,1	864	-0,1	-0,1	0,1

Note: Compiled by author.

All the information quoted above is in line with the following conclusions:

1. The opinions formed in questions X_1 and X_2 - X_{10} are mutually highly positively correlated.
2. The criterion X_2 (punctuality) is positively correlated only with two other evaluation criteria: X_1 (involvement) and X_6 (the pace of the lesson).
3. There is no correlation between the criteria X_1 - X_{10} and the opinions on the usefulness of the e-platform (X_{12} and X_{13}).
4. There is no correlation, either, between the number of opinions N and any other criterion analyzed (with the exception of the negative correlation between the number of answers and the rating of the e-platform attractiveness. It means that the bigger is the number of the people evaluating, the worse are the opinions of the attractiveness of the e-platform).
5. Negative correlation has been found between the application of the e-platform (X_{11}) and the opinions included in the questions X_6 - X_{10} . It means that the students using the e-platform usually form worse opinions in the criteria X_6 - X_{10} than the students who do not have any classes on the e-platform.

The next part of the analysis encompasses a classification of the evaluation criteria in view of their mutual correlation. The objective was to determine whether the evaluation could be limited to a smaller number of questions, given their high mutual correlation. It is a problem of redundant information and unnecessary multiplication of the same information in the subsequent criteria.

Ward's agglomeration method, with the Euclidean metric as the measure of the similarity criteria, has been applied in the classification of the ratings criteria. The calculations have been made using Statistica package. The results thus acquired have been shown in illustration 2. There is a dendrogram showing the associations of all the 11 criteria. The set of criteria analyzed here can be divided into 3 groups:

1. The first group encompasses 5 criteria (X_6 - X_{10}). The students' opinions formed in these questions are usually more critical than in the other points of the questionnaire.
2. The second group includes 3 criteria (X_3 - X_5), in which the opinions of the students are the closest to the results of the average aggregate (X).
3. The last group includes two criteria X_1 and X_2 which the students rated the most favorably when compared to the other criteria.

Generally speaking, the classification of the criteria was based on the division into the criteria with the highest ratings for the teachers (X_1 and X_2), the average ratings (X_3 , X_4 , X_5) and the relatively lower ratings (X_6 - X_7 - X_8 - X_9 - X_{10}).

The last element of the analysis are the multivariate charts using geometrical figures and graphic objects and showing several various features simultaneously (ill. 3-4). The charts include a 6-element set of the following criteria {X₁, X₃, X₄, X₆, X₈, X₉}. The charts have been created using Statistica program. Every chart has a legend which allows the readers to interpret the particular charts accordingly.

Table 6

Initial data for the correlation-taxonomic analysis

Teacher	N	X1. Teacher's involvement	X2. Punctuality	X3. Fairness	X4. Atmosphere	X5. Clarity of instructions	X. Average 1-9	X6. Pace of the lesson	X7. Way of teaching	X8. Effectiveness/efficiency	X9. Motivating	X10. Grammar vs. practical skills	X11. Total English online	X12. E-platform - attractiveness	X13 E-platform vs. efficiency
Teacher A	244	5	5	5	5	5	4,9	4,9	5	4,9	4,9	4,9	0	2	2
Teacher B	92	4,9	5	4,9	4,9	4,9	4,9	4,9	4,9	4,9	4,9	4,9	0,71	4,6	4,6
Teacher C	77	5	5	5	5	4,9	4,9	4,8	4,9	4,9	4,9	4,9	0,12	4,7	4,7
Teacher D	123	5	4,9	4,8	4,9	4,9	4,9	4,9	4,8	4,8	4,8	4,9	0,22	4,4	4,3
Teacher E	86	4,8	4,9	4,8	4,8	4,8	4,8	4,8	4,7	4,7	4,6	4,7	0,14	4,2	4,6
Teacher F	37	4,8	4,9	4,6	4,8	4,6	4,7	4,8	4,7	4,8	4,6	4,6	0,05	4	3,5
Teacher G	49	4,9	5	4,7	4,8	4,8	4,7	4,7	4,5	4,4	4,2	4,5	0,78	3,6	3,6
Teacher H	108	4,7	4,4	4,8	4,9	4,7	4,7	4,5	4,6	4,6	4,6	4,6	0,39	4,4	4,3
Teacher I	58	4,8	4,8	4,7	4,8	4,6	4,6	4,6	4,7	4,5	4,4	4,4	0,71	3,5	3,5
Teacher J	68	4,8	4,4	4,7	4,6	4,7	4,6	4,5	4,5	4,5	4,7	4,5	0,43	4,5	4,5
Teacher K	78	4,7	4,9	4,7	4,8	4,6	4,6	4,4	4,4	4,4	4,3	4,6	0,41	3,9	3,9
Teacher L	37	4,8	4,6	4,8	4,5	4,6	4,5	4,3	4,2	4,3	4,4	4,4	0,73	4,3	4,3
Teacher M	156	4,7	4,5	4,7	4,8	4,6	4,5	4,2	4,3	4,3	4,2	4,4	0,67	4,2	4,1
Teacher N	45	4,8	4,9	4,3	4,5	4,3	4,5	4,5	4,4	4,2	4,2	4,3	0,42	3,6	3,4
Teacher O	132	4,5	4,7	4,5	4,1	4,3	4,2	4,1	3,8	3,9	3,8	4	0,72	3,9	3,8

Note: Compiled by author.

Table 7

Correlation factors between the evaluation criteria included in the evaluation questionnaire

The correlation factors marked are essential when $p < .05000$ N=15 (Lack of data was removed in some cases)																	
	Av.	Dev.	N	X ₁	X ₂	X ₃	X ₄	X ₅	X	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
N	92,7	55	1	0,15	0,03	0,36	0,19	0,34	0,22	0,08	0,17	0,21	0,23	0,3	-0,28	-0,53	-0,46
X ₁ . Teacher's involvement	4,8	0,1	0,15	1	0,64	0,71	0,81	0,86	0,93	0,9	0,88	0,85	0,82	0,87	-0,46	-0,15	-0,07
X ₂ . Punctuality	4,8	0,2	0,03	0,64	1	0,19	0,31	0,39	0,53	0,65	0,47	0,45	0,29	0,46	-0,34	-0,29	-0,27
X ₃ . Fairness	4,7	0,2	0,36	0,71	0,19	1	0,79	0,92	0,8	0,55	0,69	0,77	0,79	0,81	-0,26	0,06	0,16
X ₄ . Atmosphere	4,7	0,2	0,19	0,81	0,31	0,79	1	0,85	0,91	0,77	0,92	0,89	0,85	0,93	-0,46	-0,02	0,03
X ₅ . Clarity of instructions	4,7	0,2	0,34	0,86	0,39	0,92	0,85	1	0,92	0,78	0,83	0,87	0,85	0,91	-0,36	-0,04	0,06
X. Average 1-9	4,7	0,2	0,22	0,93	0,53	0,8	0,91	0,92	1	0,92	0,97	0,98	0,94	0,98	-0,57	-0,08	-0,02
X ₆ . Pace of the lesson	4,6	0,3	0,08	0,9	0,65	0,55	0,77	0,78	0,92	1	0,92	0,89	0,82	0,87	-0,56	-0,18	-0,14
X ₇ . Way of teaching	4,6	0,3	0,17	0,88	0,47	0,69	0,92	0,83	0,97	0,92	1	0,95	0,91	0,94	-0,59	-0,12	-0,07
X ₈ . Effectiveness/efficiency	4,5	0,3	0,21	0,85	0,45	0,77	0,89	0,87	0,98	0,89	0,95	1	0,96	0,97	-0,64	-0,01	0,03
X ₉ . Motivating	4,5	0,3	0,23	0,82	0,29	0,79	0,85	0,85	0,94	0,82	0,91	0,96	1	0,95	-0,62	0,05	0,09
X ₁₀ . Grammar vs. practical skills	4,6	0,3	0,3	0,87	0,46	0,81	0,93	0,91	0,98	0,87	0,94	0,97	0,95	1	-0,61	-0,04	0,02
X ₁₁ . Total English online	0,4	0,3	-0,28	-0,46	-0,34	-0,26	-0,46	-0,36	-0,57	-0,56	-0,59	-0,64	-0,62	-0,61	1	0,2	0,2
X ₁₂ . E-platform – attractiveness	4,0	0,7	-0,53	-0,15	-0,29	0,06	-0,02	-0,04	-0,08	-0,18	-0,12	-0,01	0,05	-0,04	0,2	1	0,96
X ₁₃ . E-platform vs. efficiency	3,9	0,7	-0,46	-0,07	-0,27	0,16	0,03	0,06	-0,02	-0,14	-0,07	0,03	0,09	0,02	0,2	0,96	1

Note: Compiled by author.

Table 8

Levels of significance of the correlation factors between the evaluation criteria included in the evaluation questionnaire

	N	X ₁	X ₂	X ₃	X ₄	X ₅	X	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
N		0,6	0,93	0,19	0,49	0,21	0,43	0,78	0,53	0,45	0,42	0,28	0,32	0,04	0,08
X ₁ . Teacher's involvement	0,6		0,01	0	0	0	0	0	0	0	0	0	0,09	0,59	0,79
X ₂ . Punctuality	0,93	0,01		0,5	0,26	0,15	0,04	0,01	0,08	0,09	0,29	0,09	0,21	0,29	0,33
X ₃ . Fairness	0,19	0	0,5		0	0	0	0,04	0	0	0	0	0,35	0,83	0,58
X ₄ . Atmosphere	0,49	0	0,26	0		0	0	0	0	0	0	0	0,08	0,94	0,91
X ₅ . Clarity of instructions	0,21	0	0,15	0	0		0	0	0	0	0	0	0,19	0,89	0,83
X. Average 1-9	0,43	0	0,04	0	0	0		0	0	0	0	0	0,03	0,77	0,94
X ₆ . Pace of the lesson	0,78	0	0,01	0,04	0	0	0		0	0	0	0	0,03	0,51	0,61
X ₇ . Way of teaching	0,53	0	0,08	0	0	0	0	0		0	0	0	0,02	0,66	0,82
X ₈ . Effectiveness/efficiency	0,45	0	0,09	0	0	0	0	0	0		0	0	0,01	0,97	0,93
X ₉ . Motivating	0,42	0	0,29	0	0	0	0	0	0	0		0	0,01	0,87	0,75
X ₁₀ . Grammar vs. practical skills	0,28	0	0,09	0	0	0	0	0	0	0	0		0,02	0,9	0,95
X ₁₁ . Total English online	0,32	0,09	0,21	0,35	0,08	0,19	0,03	0,03	0,02	0,01	0,01	0,02		0,47	0,47
X ₁₂ . E-platform – attractiveness	0,04	0,59	0,29	0,83	0,94	0,89	0,77	0,51	0,66	0,97	0,87	0,9	0,47		0
X ₁₃ . E-platform vs. efficiency	0,08	0,79	0,33	0,58	0,91	0,83	0,94	0,61	0,82	0,93	0,75	0,95	0,47	0	

Note: Compiled by author.

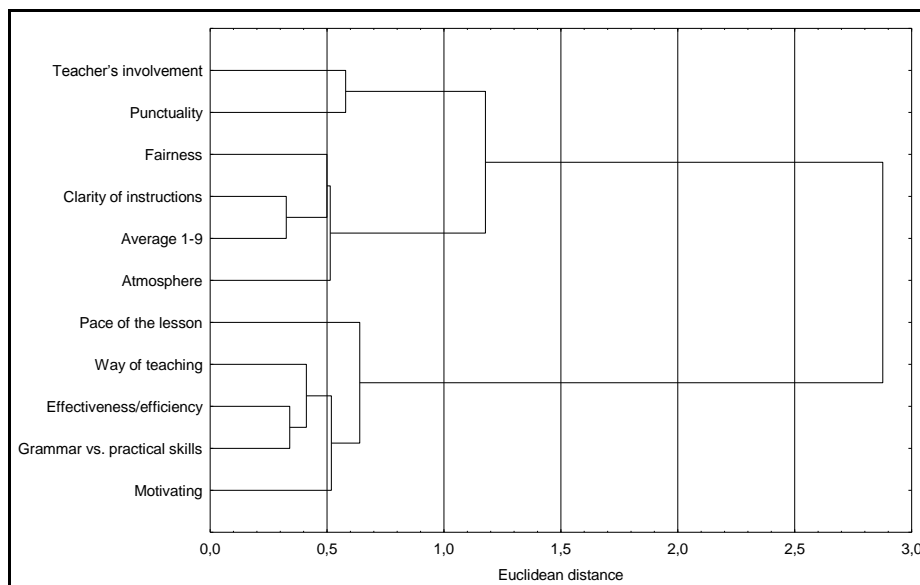


Figure 2. Ward's dendrograms of the evaluation criteria.

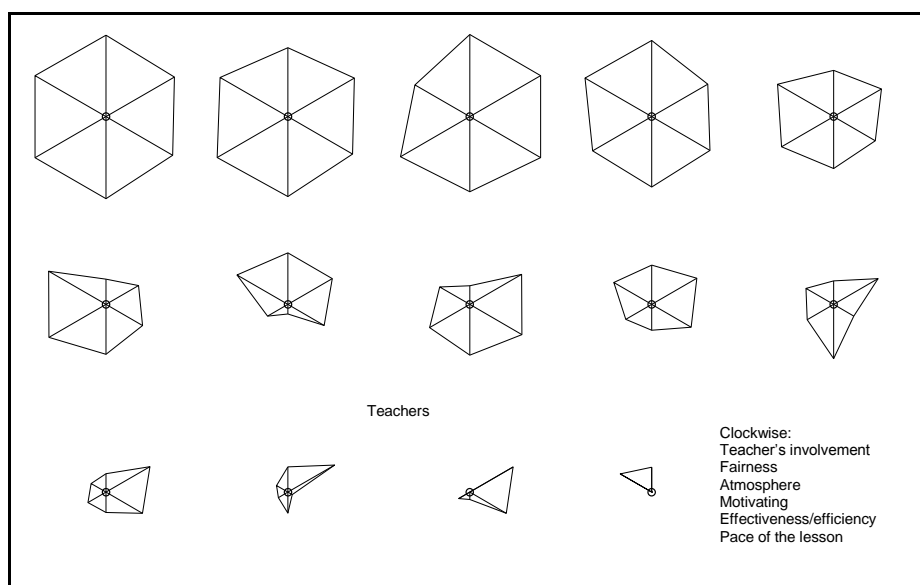


Figure 3. 6-dimensional profiles of the teachers.

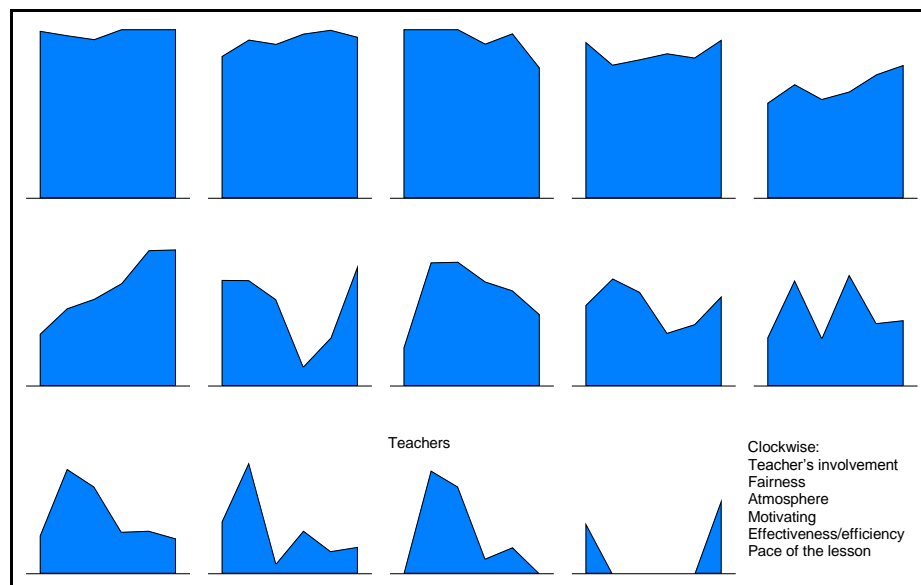


Figure 4. Another 6-dimensional profiles of the teachers.

5. ANALYSIS OF THE TEXT OPINIONS

Text opinions, which the students include in the open question (appendix A), are an interesting element of the evaluation. Altogether there were 281 texts – 210 texts with different wording and 71 identical ones, e.g. *no comments*, from a single word to a few sentences. There is a reference to the particular teacher that is being evaluated next to every remark.

The remarks have been grouped into 8 categories:

1. EXCEL = EXCELLENT – a very favorable text (20)
2. GOOD – a comment that includes positive remarks (109)
3. NONE – a text meaning that the student has no remarks to make (9)
4. !!! – a remark negative in meaning (2)
5. MET = METHODS – the usefulness of the methods of teaching used in class (17)
6. BOOK – a remark or suggestion connected with the coursebook (7)
7. CLIP – a comment on the usefulness of the e-workbook (15)
8. ORG – suggestions on how to modify the lesson (31)

On the basis of the numbers of the remarks, quoted in brackets, the following conclusions can be drawn:

1. Very good and good remarks prevail – there are 47% of them (131 out of 281), or 60% (129 out of 210) when identical remarks are eliminated from the list.
2. Negative opinions are scarce (only 2) and they constitute less than 1% of the remarks.
3. The remarks on the organization of the lesson constitute a large portion (over 30, i.e. 12/15%) of all.
4. Relatively few remarks (7) were made in the section PODR (coursebooks).

6. FINAL REMARKS

The results of the questionnaire allow one to draw a conclusion that language courses at the Andrzej Frycz Modrzewski Krakow University are taught in a professional way and at a high level. The teachers are involved, they explain the tasks set and the issues presented in a clear way, the classes are run in a friendly atmosphere, at a right pace and in an interesting style, and – most importantly – they are efficient. The teachers are trying to motivate the students the best they can and they fairly evaluate their workload and achievements. On a scale from 1 to 5, the average rating of the language courses varies from 4 to 5, which is a very high result, especially when seen in the context of the evaluation process of other classes, the basic and the professional ones.

Apart from the sum of points for the answers to ten basic questions evaluating the teacher's work, there have been two additional questions given that referred to the teaching using the electronic form of the workbook and one open question that allowed the students to express their general views on the language courses and the teachers running these classes.

In the last part of the questionnaire special attention should be given not only to the very positive and often repeated opinions about the teachers, but also to the remarks about the teaching process that would allow the management of the Foreign Languages Center at the Andrzej Frycz Modrzewski Krakow University to improve it.

Several critical opinions have been noted, as well, and they referred to the excessive number of students in one group, placing students at various language levels in the same group, late hours of the language classes, the obligatory learning materials required by the Foreign Languages Center (SJO) that must be purchased by the students and some problems in using the educational platform. Many of the problems listed must be discussed with the University Authorities, who could give their permission to innovate the organization of the language courses, especially in the master studies.

When interpreting the results the following issues must be considered:

- 1) It is the first evaluation using the new formula. Therefore, it has been assumed that its results cannot be compared with the former evaluations. For this reason this elaboration is an analysis of one examination only, with no reference to the former evaluations.
- 2) It follows from such an assumption that the teachers' ratings should be interpreted carefully. It will be possible to consider their diagnostic value only if these ratings turn out to be similar to the next 2-3 evaluations.
- 3) The number of answers given must be considered, too. They vary from 37 to 244. Therefore, the average results have different diagnostic values.
- 4) Because of the limited volume of this report, the deviations from the standard ratings have been neglected. Hence, the trust intervals for the average ratings have not been determined, which could improve their diagnostic value.
- 5) Another important problem to consider is the objectivity of the evaluation ratings in the context of the lecturers' requirements for their subjects. In order to determine if there is such a link, a correlation analysis between the lecturers' ratings and the final grades they give should be carried out. Such analysis is planned after the next edition of the evaluation.
- 6) Every language teacher receives a complete report of the evaluation in which the surnames are substituted with the teachers' codes (A, B, C,...). The teacher is informed about his/her code but doesn't know the codes of the others. This way every teacher has all the information about his/her evaluation results, knows his/her place in the rating which is determined on the basis of the average ratings given by the students, but at the same time full anonymity of the evaluation results is ensured. If the rating falls below the teacher's expectations, he/she can identify the faculty, the year and the form of studies (full-time or part-time) and then consider the possible reasons for the low results, his/her mistakes and the possible solutions and ways to avoid them in the future.

REFERENCES

- Millman, J., & Darling-Hammond, L. (1990). *The New Handbook of Teacher Evaluation: Assessing Elementary and Secondary School Teachers*. Corwin Press, Inc., A Sage Publications Company, 2455 Teller Road, Thousand Oaks, CA 91320 (paperback: ISBN-0-8039-4523-X).
- Patton, M. Q. (1987). *How to use qualitative methods in evaluation* (Vol. 4). Sage.

Appendix A.

A sample list of text opinions included in the open question

EXCEL	Being a middle-aged student, I am very pleased with the progress and the form of the classes. Especially for a person that has never learned English before. The involvement of the teacher is very helpful and it affects my problems with English in a positive way. I hope that this teacher will stay with our group in our third year.
EXCEL	The English classes are interesting, which makes learning much easier.
EXCEL	It is my wishful thinking – if there is such a possibility, I really wish that – that we have the same English teacher in our Master Studies.
EXCEL	The best teacher at the university.
EXCEL	No comments. Learning English with this teacher is a pleasure. Attractive classes and a large involvement of the teacher in my development.
EXCEL	The professor teaches in an interesting, motivating way. The atmosphere is so positive that it even motivates the students to learn other subjects. In my opinion, this teacher sets an example for others.
EXCEL	This is the best teacher I've met at the Krakow University.
EXCEL	I would like this teacher to continue teaching our group in our Master Studies. She has a big heart, she is exceptionally patient, understanding, warm and respectful for the students! Thank you in advance.
GOOD	Can explain things well, a great teacher.
GOOD	A nice and friendly teacher; may we have more such teachers.
GOOD	A very positive teacher, fair when grading the students.
GOOD	We would like this teacher to continue teaching our group.
GOOD	We would like this teacher to return to teaching our group.
GOOD	Very high level of classes.
GOOD	I want to have classes with this teacher. Great approach, involvement and motivation.
GOOD	I am very pleased with my English classes. I wish we had more.
GOOD	The lessons do not require any modification. They are transparent and allow everyone to understand the issues discussed in class.
GOOD	The classes conducted in a friendly way, the tests are clear and only include the material discussed.
GOOD	The teacher is very fair. She has vast knowledge and she knows how to pass it on. The only thing I wish is that we had more English classes.
GOOD	I would like this teacher to continue teaching our group. I am very pleased with the way of teaching because she can pass on the right knowledge in a clear and simple way.
NONE	No comments on the rationalization of the English classes at the KAAFM.
NONE	I have no comments or suggestions.
NONE	I found it hard to answer some of the questions because we had few classes with this teacher so far.

[!!!]	The amount of material and hours is catastrophic, too small per one semester. The same applies to small tests that would motivate students to learn English systematically and would allow them to reach a higher level of proficiency. The lack of the division into groups is a complete failure. How is it possible that the people who have never learned English are in one group with the people who went to an extended English course in high school? The prices of the coursebooks and the access cards are much too high, the cards should be given free of charge because the activity on the platform is a requirement. The platform itself includes many mistakes.
[!!!]	The teacher can't explain grammar. It seems to me she doesn't know the basic notions. I haven't learned anything new yet.
MET	We should have more conversations.
MET	The pace of the information passed in class could be higher. The time of the lesson is used to a satisfactory extent.
MET	In the case of classes conducted in English, new notions and issues should be introduced in Polish to give student more psychological comfort. The pace of the lesson and the level of the class are too high since the beginning of the semester, at this stage of learning.
MET	More Polish translation.
MET	More focus on conversation.
MET	More interesting vocabulary items, less grammar!
MET	More listening – films, „real life” situations, to minimize the element of surprise that a student gets abroad when he encounters a different English, spoken faster.
MET	More legal vocabulary.
MET	More classes connected with business English and management.
MET	More focus on colloquial speech!
MET	The teacher should focus on the pronunciation of English words and should check the students, otherwise the students learn to mispronounce. I think that the overall grade cannot be lowered because of the absences. Learning a language is about the level of knowledge and not the number of absences. We can't have English classes on Friday evenings till 8.45 p.m. or on Saturdays because we are on full-time studies.
MET	Too much grammar and too little CLIP in the overall grade.
BOOK	Give us a possibility to buy used books, this is thievery!
BOOK	Using books that cost 100zl is irrational. Moreover, the content is not worth such a price. At least the one that we are using now.
BOOK	The learning materials required by the university are, to put it mildly, worthless and uninteresting, making teaching difficult. Not dividing the students into groups according to their level of language is a big mistake. The system of teaching English is too expensive and troublesome in the formal sense. In this sense, we don't have time for that! And your keyboard is broken, there is no “a with a diacritic mark” key. Please change something, and make it quick!
BOOK	In my opinion the books are too expensive. They should be much cheaper.

BOOK	I don't like the policy of the book purchase. The fact that we can buy the book only in the university bookstore and at such a high price is absurd. As for the classes, we should have more grammar and faster pace of working with the book (instead of focusing on every instruction, it's better to cover a larger portion of the material).
BOOK	Too expensive books. Only buying the book with the clip access card guarantees a positive overall grade – extortion! Apart from that, the classes are impeccable.
CLIP	I would like to have an opportunity to train using video.
CLIP	I think clip is a mistake.
CLIP	The platform makes no sense, a useless expenditure.
CLIP	The platform isn't adjusted to the classes.
CLIP	It's necessary that we could watch films in Polish.
CLIP	The CLIP system is faulty, in my opinion. It happened many times that I entered a correct answer and yet the system marked it as incorrect. To my surprise, when I clicked on the option "show the correct answers", my previous answer was listed there, too. Although CLIP marks the answers in accordance with an answer key, it is inaccurate in that. Long Live the King!
CLIP	More clip.
CLIP	Turn off clip!
CLIP	The tasks on clip should be closer to the material we cover in class. Besides, we should practice grammar and vocabulary more.
ORG	English classes are too late. In my opinion, they should be in the morning.
ORG	Office hours for part-time students – in the afternoon.
ORG	Awful time schedule of our English classes!
ORG	The main factor that lowers the attractiveness of the classes is the lack of division into groups according to the students' level and abilities. The people who do not have the basic knowledge of English lower the level of the group for the people who have much better qualifications. I suggest that the lessons focused more on conversation in the second language, rather than on solving problems or doing the exercises from the book.
ORG	If there is a possibility to form several language groups, it is definitely advisable. The differences in level of English among the students in our group are so big that it often renders successful cooperation impossible.
ORG	Smaller groups and placement tests to divide students into groups according to their knowledge.
ORG	The division into groups according to the students' level is the most suitable form of teaching. In our faculty there was only one group because of the small number of students but the teacher was trying to find an appropriate task for every student in class.
ORG	Don't change the teacher in the course of the semester.
ORG	Watching films in English should be introduced.
ORG	The schedule is abominable, which isn't the teacher's fault, but the dean's office's that decides about all the schedules. On many occasions the classes were rescheduled, cancelled or overlapped with some other classes, partially

	or completely. It isn't a problem of the English course only, the same applies to all other classes, too. The teachers are unhappy with this situation, as well, and getting from Lipowa Street to Building A or B in a 5-minute break IS IMPOSSIBLE!
ORG	In general, English is a cool subject but scheduling the classes for Friday, late in the evening, or for Saturday, for full-time students, who spend every day at the university, from dawn to dusk, is just inappropriate. It is insane, I hope that the Department of Education will understand their mistake.
ORG	First thing: there should be a division into groups according to the students' level of language, we should have more practical preparation for our profession, discuss more.
ORG	My first remark about teaching English: the way of correcting the exercises sent by the clip system is not very reliable because even if the whole task is grammatically correct, points are still deducted for not putting a full stop or using too many spaces between words. There is no possibility to correct the task afterwards. I think that the clip access cards should be given free of charge because doing the tasks is obligatory. Another points is about the way of teaching. I have no objections here but I think that the lessons should motivate students more to learn individually.
ORG	A division into groups according to the students' level of English. Slower pace of the lesson that would allow a deeper analysis of the material and the tasks discussed; giving the key words / a vocabulary list after every lesson to be learned by the students; introducing revisions before tests and exams that facilitate passing them.
ORG	I would like to see the end of the monopoly on the English books. I think that there should be a possibility to pass the English course in a form of an additional final exam or something like that for the students on a high language level. Attending an English course on a low level discourages students and is a waste of their precious time which they could otherwise devote to attending a language course on a higher level or just use more effectively.

Primljeno: 05.10.2013.

Odobreno: 18.10.2013.