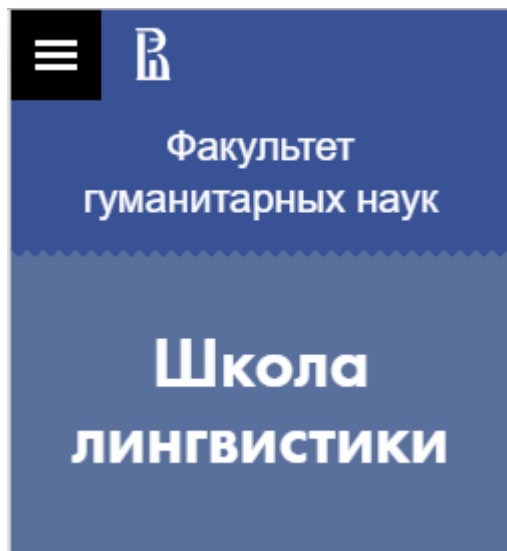




Пути использования учебных корпусов и компьютерных инструментов в самостоятельной работе изучающих иностранный язык и в практике преподавания иностранного языка



Виноградова Ольга  
Ильинична, доцент школы  
лингвистики НИУ ВШЭ  
[olgavinogr@gmail.com](mailto:olgavinogr@gmail.com)

# Компьютерные инструменты ШЛ

- Автопомощник в определении ошибок в письменных текстах ADWISER (Automated Detection of Writing Inaccuracies for Students of English in Russia).
- Автоматизированная оценка сложности текста с выводением примерной оценки за эссе одного из двух экзаменационных жанров - приложение INSPECTOR
- Платформа для автоматического порождения тестовых вопросов на основе ошибок, аннотированных специально подготовленными студентами, и исправлений этих ошибок - RETM=REALEC Test Maker.

## Использование компьютерных инструментов в самостоятельной подготовке к экзамену по английскому языку

- Загрузка написанного текста в ADWISER, получение выделенных областей в тексте, которые ADWISER счел содержащими ошибку, исправление тех из выделенных областей, в которых автор видит ошибку после указания на нее; проведение повторных проверок до тех пор, пока есть что исправлять.
- Загрузка полного текста эссе в INSPECTOR, получение автофидбека по 59 показателям сложности текста наряду с приблизительной оценкой по параметрам сложности; возможное усовершенствование текста и повторное проведение оценки его сложностей.

# Использование компьютерных инструментов преподавателем EFL

- Преподаватель загружает текст студента в ADWISER, получает выделенные области в тексте, которые ADWISER счел содержащими ошибку; по желанию добавляет еще выделения областей ошибок; загружает текст в учебный корпус REALEC; преподаватель или студент размечает ошибки тэгами из предлагаемой классификации и предлагает исправление области ошибки; преподаватель нажатием кнопки Create quiz получает заготовку для теста по допущенным ошибкам, редактирует его и дает студенту после работы студента с ошибками; преподаватель или студент загружает текст в INSPECTOR, получает автофидбек по 59 показателям сложности текста и оценку его сложности студенту; в результате оценка за работу с текстом может включать оценки каждого этапа из описанных выше.

# Работа по усовершенствованию автопомощника ADWISER

Стадии работы:

➤ Создание баз данных подозрительных контекстов. Примеры контекстов:

использование в одном предложении формы Present Perfect и предлога in со следующим за ним числительным; та же форма глагола и последовательность from+числительное ... to +числительное; та же форма +between + числительное + and + числительное; Wh-word + NP + VP (+... )? (\*Why you are here?)

➤ Анализ баз данных с последующим дописыванием ограничений на вхождение в число подозрительных контекстов.

➤ Получение статистики по каждому контексту на больших массивах.

На данном этапе: Present Perfect, Past Continuous, порядок слов в прямых и косвенных вопросах, условные предложения, запятые, согласование подлежащего и сказуемого и существительного с указательными местоимениями

TreeTagger + Python reg ex

## Примеры для базы данных

Summing up, the population of the USA has been increasing steadily throughout the history and reached the number of 300 million people in the beginning of 21 century

3

Moreover, in Latin America in 2014: unemployment rate has become lower than in 2014

0

The main information these graphs give is that, firstly, more people have started to use mobile Facebook application instead of the desktop one and, secondly,  
The first bar chart shows that the number of people using desktop Facebook version in 2013 is slightly smaller than in 2012: approximately 137 millions of users versus 150 millions

0

To sum up: LinkedIn has gained more users than Instagram in every strata, except for the 18-29, where Instagram has % of users now that doubled in comparison with LinkedIn

5

For example, in 1999 there was a massive homicide by two young guys who have killed a plenty of innocent students (their mates also)

0

The one of the most popular question in the world in the 21 century is how modernization in technologies has infuenced on every person

1

# Типичные для русскоязычных авторов ошибки L1 Interference

Инверсия в косвенных вопросах: *They should control what their child watches, reads or plays, and who **are her companions**.*

Отсутствие инверсии в прямых вопросах: *Why **this is** happening in our instantly developing world?*

Перемена мест подлежащего и именной части сказуемого: ***Happy** were **only small children**.*

Вынесение прямого дополнения в начало: *As we see it on one of the charts in 1999 **France** have visited **about 11 millions of UK visitors**. On the other hand , **the minimum of proportion the countries had** in 1940, which was about 5-10%. **The same** we can say for 50-64 year group of users.*

Вынесение обстоятельства перед сказуемым: *...**how the percentage of mothers of many children ... so much has fallen**(,) that ...*

Неправильная позиция обстоятельств, связанных с выбором видо-временной формы глагола: *If **universities ever** would count gender as a decisive factor, the whole education system would be broken in my opinion.*

Неправильная позиция слова *also*: *We **also can** notice that dead rate is low.*

Неправильная позиция других обстоятельств и парентетических элементов: *...**the amount of people to every discipline by no means** should be determined by universities. Increased number of, **for example**, gyms will contribute to the easier accessibility to such sport facilities.*

Вынесение элемента, оказавшегося в фокусе, в начало - иногда с инверсией: *For example, **in the USSR** was the ideal situation.*



# Пробная работа с ADWISER

<http://34.65.103.57:5000/>

Скопировать текст в окошко и нажать кнопку “Annotate”.

Если обнаружены ошибки, они будут выделены в тексте красным фоном. Нажмите на зону ошибки, и вы увидите сообщение с текстом о характере ошибки. При повторном нажатии сообщение исчезнет. Для того, чтобы вернуться назад и определить ошибки в новом тексте, нажмите Back.

# Работа с системой оценки сложности текста Inspector

- [https://linghub.ru/inspector\\_realec/](https://linghub.ru/inspector_realec/)

### Inspector

#### Essay

The following graph shows approximate time people in England in 2012 spent on sport exercises. It gives figures for both men and women and it can be clearly seen that people in different age categories spent different amount of time doing sport. According to this diagram, men of 16-24 age sport 282,1 minutes on exercises and it is the highest time among age and gender categories. However, women of 16-24 spent 114,1 minutes, which is less than men in the same age group but it is the highest rate for all women's age groups. Average time for all men and women is nearly 114 for men and almost 93 minutes for women. Furthermore, the smallest amount of time for both genders was indicated in the group of people who are almost 75 years older. In this category men spent 29,2 minutes and women spent 10 minutes. It is also notable that average sport time of women who are 25-34 and 35-44 years old did not change and remained at 42,8 minutes.

файл не выбран

Inspect

#### Result

Your essay is not good enough. The probable grade in the exam is **5**. To improve your writing skills look at the recommendations below.

#### Statistics

Number of words	177
Number of lemmas	88
Number of sentences	8

#### Information about academic words and linking phrases

Number of academic words	6
Number of linking phrases	11

The **following graph** shows **approximate** time people in England in 2012 spent on sport exercises. It gives figures for both men **and** women **and** it can be clearly seen that people in different age **categories** spent different amount of time doing sport. According to this **diagram**, men of 16-24 age sport 282,1 minutes on exercises **and** it is the highest time among age **and** gender **categories**. However, women of 16-24 spent 114,1 minutes, which is less than men in the same age group but it is the highest rate for all women's age groups. Average time for all men **and** women is nearly 114 for men **and** almost 93 minutes for women. Furthermore, the smallest amount of time for both genders was indicated in the group of people who are almost 75 years older. In this **category** men spent 29,2 minutes **and** women spent 10 minutes. It is **also** notable that average sport time of women who are 25-34 **and** 35-44 years old did not change **and** remained at 42,8 minutes.

#### Spelling mistakes

It is also **notible** → **notable** that average sport time of women who are 25-34 and 35-44 years old did not change and remained at 42,8 minutes.

## Recommendations

The model that predicted your grade is based on the features presented in the table below. Having compared the mean values of these features, we highlight the features that you should improve (**green** - good indicators; **red** - features to be improved). All in all, in your essay you should use more different verbs; linking phrases; gerunds; morphologically complex words; infinitives; different words; functional n-grams; complex sentences; relative clauses; words before the main predicate; possessive constructions. You used enough long sentences; academic words; different nouns; verbs in past simple; auxiliary verbs; sophisticated verbs; nouns, verbs, adverbs, and adjectives; coordinate constructions.

Your essay	Feature	Best
9	Number of coordinate phrases	6
0.47	Lexical density	0
0.59	Verb sophistication	0.31
0.85	Number of auxiliaries	0.61
6	Number of verbs in past simple	4
0.1	Noun variation	0
0.22	Lexical sophistication	0.12
24	Average length of the sentence	23
2.16	Verb variation	2.82
13	Number of linking phrases	14
1	Number of gerunds	3
0.03	Derivational level 3	0.05
0.0	Derivational level 4	0.07
2	Number of infinitives	11
88	Number of lemmas	131
0.0	Derivational level 6	0.04
0	Number of functional n-grams	1
17	Number of clauses	25
1	Number of adjective clause	5
5.12	Average number before root	6.4
8	Number of sentences	13
0.05	Derivational level 5	0.06
15	Number of possessives	16

# 59 параметров сложности текста в системе Inspector

[https://docs.google.com/spreadsheets/d/14C0HqPtCSBb\\_uPRGj3A1HsVlQypsdIyaZfOs\\_\\_OKGI4/edit#gid=0](https://docs.google.com/spreadsheets/d/14C0HqPtCSBb_uPRGj3A1HsVlQypsdIyaZfOs__OKGI4/edit#gid=0)

# Russian Error-Annotated Learner English Corpus

<http://realec.org/>

HOME

Texts ▾

Browse

Download

About

Contacts

## Russian Error-Annotated Learner English Corpus



Total 14361 texts/3603116 words

Exam 5965 texts/1412169 words

Total 13585 annotated texts

### Acknowledgements

To Prof. Natalia Kolyadina for providing student texts from the Independent Examination in English



Russian Error-Annotated Learner English Corpus by REALEC Team

is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

[Permissions beyond the scope of this license may be available here.](#)



HIGHER SCHOOL OF ECONOMICS  
NATIONAL RESEARCH UNIVERSITY

# Схема работы в корпусе REALEC - 1

- Выбрать вверху на заглавной странице корпуса кнопку Texts
- Выбрать папку exam
- Выбрать в оглавлении папки раздел exam 2018
- Войти в первый текст
- login & password -на всплывающей кнопке в правом верхнем углу ввести одну из следующих пар логина и пароля:

Cseminar 1234abc

Cseminar1 1234ABC

PSeminar1 1234abc

PSeminar2 1234abc

PSeminar3 1234abc

PSeminar4 1234abc

PSeminar5 1234abc

# Схема работы в корпусе REALEC - 2

Для загрузки текста зарегистрировавшемуся пользователю надо сделать следующее:

- Кнопка DATA, всплывающая в левом верхнем углу
- Выбрать окно Import
- В окне TextID дать название вашему файлу латиницей
- В окне Text скопировать сохраненный текст
- Подтвердить нажатием ОК
- Для открытия окна аннотации выделить область ошибки

Можно таким же образом загрузить текст со следующего слайда.

# Пробный текст.

For the next ten years in Sweden the quantity of the old people have been falling and then, in 1990, started to rise to 20% in 2010. The tendency of developing countries is different: from 2006 to 2008 the number was increasing, in 2009 the number dropped, from 2010 to 2011 it was growing again and from 2011 to 2013 the amount was decreasing. We can see, when the data of the USA and Sweden has grown up to 15%, the 65 years old population in Japan was levelling out to 3% and stayed approximately the same from 1960 to 1990. Another interesting fact is, that at the beginning the population aged 65 was less in China and the most in USA, but in 2040 the situation changed completely and the population aged 65 was the biggest in Japan, Sweden has always stayed in the middle. If we will believe this forecast, then all the countries will have more elderly people. Never in their life people have been living longer than nowadays.



# Тестмейкер RETM

1. Для того, чтобы создать небольшой квиз на материале аннотированного эссе, нужно навести курсор на верхнюю панель и выбрать кнопку “Create test”



# Тестмейкер RETM - 2

В появившемся поле введите логин и пароль пользователя тестовой платформы (требуется сделать 1 раз, далее логин и пароль будут сохранены и заново их вводить не потребуется) - логин и пароль выдаются по требованию:

Please login as a teacher or admin to continue

Your username:

Your password:

[Log In](#)

## Тестмейкер RETM - 3

После этого отобразятся поля со сгенерированными вопросами и ответами. Область ошибки определяется текстом, выделенным жирным шрифтом, при изменении выделенной жирным области, изменится и область ошибки. В этом режиме доступно только базовое редактирование вопросов, расширенные возможности по редактированию вопросов (удаление правильного ответа и добавление дополнительного правильного ответа). Для того, чтобы создать новый квиз, прокрутите страницу до конца, введите в поле внизу название квиза, отметьте галочкой поле “Generate new quiz” и нажмите кнопку “Save questions”

После этого у вас отобразится список доступных вам в системе квизов. Для того, чтобы вернуться к сгенерированному ранее квизу, нажмите кнопку “Edit” рядом с его названием.

# Тестмейкер RETM - 4

После этого у вас откроется с расширенным редактированием выбранного квиза

- [Questions](#)
- [Users](#)
- [Log Out](#)
- You can edit questions and right answers in text areas here
- To delete question from quiz tick the box "Delete this question from quiz"
- To delete answer from the set of right answers tick the box "Delete this answer from question"
- To add one more right answer to question click "Add another right answer" after last answer
- If you want to discard changes you've made to a question or an answer click "Restore to default"
- You can select whether answers submitted for this quiz should be stripped from punctuation
- When you're done click "Save changes". You won't be able to discard changes you've made after that.

## Quiz Example quiz

[Scroll to bottom](#)

Maximum questions per page:  [Change](#)

Page 1 out of 1

Pages: [1](#)

### Question 0

Nowadays, we have buildings of all shapes and colors, whatever their actual purpose is. Still, many people think that we should not pay so much attention to what the building looks like, but be more concerned about **the reason it for what case was built**. I disagree with this opinion.

Delete this question from quiz [Restore to default](#)

### Answers

the reason it was built for

Delete this answer from question [Restore to default](#)

# Тестмейкер RETM - 5

Если отметить галочкой поле перед текстом, начинающимся со слова “Delete”, соответствующий вопрос/правильный ответ будет удалён из базы данных после нажатия кнопки “Save changes”

- После нажатия кнопки “Restore to default” текст в поле вопроса/правильного ответа вернётся в изначальное состояние (до внесения несохранённых изменений)

- После нажатия кнопки “Add another right answer” под первым правильным ответом появится поле для ввода ещё одного правильного ответа на вопрос

- После нажатия кнопки “Save changes” все введённые вами изменения сохранятся без возможности отмены

- Чтобы изменить количество вопросов, отображаемое на одной странице, введите нужное вам количество вопросов в поле “Maximum questions per page” и нажмите кнопку “Change” рядом с ним.

- Для того, чтобы быстро перемещаться между верхней и нижней частью страницы используйте кнопки “Scroll to bottom” и “Scroll to top”

7. Также вы можете сами пройти созданный вами тест. Для этого вернитесь на страницу со списком тестов и выберите “Preview” рядом с названием квиза. Отобразившаяся после этого страница будет соответствовать тому, что будут видеть студенты при прохождении квиза.

# Частотность разных групп тэгов (по материалам одной папки)

Lexical choices	14607
Spelling	11829
Discourse-organising	10498
Articles	10323
Punctuation	6770
Verb forms	6061
Prepositions	4310
Choice between Sg/Pl for nouns	2594
Word order	1547

# Correlation of work in the corpus with success in the English examination

		All students	Students working with REALEC
2014-2015	2nd year	7.43	8.81
2014-2015	1st year	7.33	7.95
2015-2016	2nd year	7.45	7.75
2015-2016	1st year	6.84	7.05(min 4,5)
2016-2017	2nd year	7.5	7.7(min 7,7)
2016-2017	1st year	7.33	7.85
2017-2018	2nd year	7.67	7.71