

Creation Technology of Electronic Distance Course

Nikolay Karpov¹

¹National Research University Higher School of Economics, Nizhny Novgorod, Russia
nkarpov@hse.ru

Abstract. Technology of electronic distance course development was created. This technology is useful for developing training tutorial of foreign languages. Firstly, it can be integrated into the LMS and can be used online as a web service. Secondly, it can be used as a standalone desktop tutorial. A helpful tool tip was added. The tip contains grammar information and English translation for each separated word in the tutorial. With the help of this technology there was created a distance course for Russian language as a foreign one. Additionally, on the basis of Russian National Corpus there was organized a tool for search and extraction of actual sentences with examples of word usage.

Keywords: multimedia information systems, distance learning, foreign language, semantic modeling.

1 Introduction

Since the end of 19th century in Russia appeared some forms of distance learning. Russia started to use this form more frequently after the wide dissemination of open Internet technology. The government in Russia legalized the form of distance education, in addition to a full-time, a part-time and a correspondence education in February 2012 [1]. Electronic training materials became more popular than traditional textbooks. With some specific conditions it can not only increase training efficiency, but convert it to the new format.

The quality of distance learning substantially depends on using technology. Complexity of using electronic tutorial can not only hamper the perception of teaching materials, but it can cause aversion to the tutorial. Thus, the successful implementation of e-learning is based on the choice of the software appropriate for the specific requirements, the goals and objectives set to it. Despite the enormous number of existing methods of data processing, their subsequent layout and transfer of learners, not all of the methods have the desired characteristics. In addition, linguistic electronic course should have its special characteristics.

The purpose of the work presented is technology development of creation of remote education course. We want to support two ways to use our course. First way, we need to use this tutorial as a part of web resource. Second way, we want to have the ability to use it locally on your personal computer as a standalone version.

We developed the technology for tutorial creation as a linked HTML (HyperText Markup Language) files which support the format of the SCORM (Sharable Content

Object Reference Model) using eXeLearning system [2]. The developed technology is used to compose the material of the tutorial on Russian as a foreign language named «The Use of Verbs with Prefixes». The special feature of this electronic course is the possibility to view the sentences from RNC (Russian National Corpus) [3] in real time, which is actual examples of using verbs with prefixes. Also we create possibilities to use a tool tip containing the word grammar information, as well as lemma (basic word form) and word translation to the English language.

2 The Tutorial Structure

As a platform for the filling of the electronic distance course in the format of SCORM course was chosen XHTML (Extensible Hypertext Markup Language) open source editor of materials for e-learning – eXeLearning. This editor is a tool for teachers and scientists, allowing you to design, develop and prepare for the publication of educational and methodological materials in electronic form.

In the process of the tutorial development there was solved a number of problems. Correspondence to the structure of educational material was developed by a tree-like structure of the electronic resource. Using the editor eXeLearning organized theoretical and practical part of the book, where were added basic information elements.

The theoretical part of the tutorial is filled with the pre-selected materials from RNC. The main component of the theoretical part of the book is the rules of the use of the prefix to the verb and a number of examples, which are available after pressing the corresponding button.

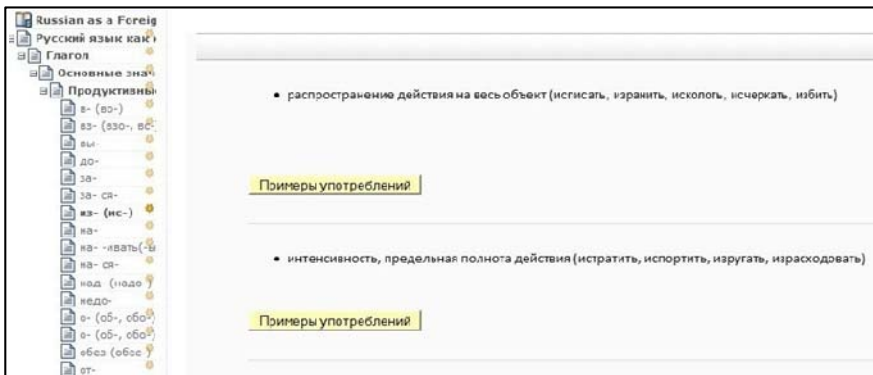


Fig. 1. Tree-type structure of the tutorial.

- a non-word spelling sign of Russian written language. According to other terminology is one of the superscript of diacritical marks, which puts over a vowel and corresponding impact sound.

, which is encoded as U+0301 (769 in the decimal system, that is in the text of an html

document should write ́). This sign and others like this one, combining characters, if there is an after percussion letters then combined with it. That is, emphasis will be placed above the letter, and not after it [4].

The advantage of this method is that we can put the accent mark over any letter. The program, which can't correctly display the symbol, also understood that it is a complementary character and simply skips it. It is important for us in the word automatic translation and in generating grammar information.

The practical part of the electronic tutorial is called «The Use of Verbs Prefixal in the Speech. The exercises». In this section, we added a variety of exercises for single choice, multiple choices and filling the gaps. Developed special video activities, for which the most appropriate movies from the media corps RNC was selected. All videos are hosted on the video-hosting service «Yandex Video», and into the exercise were integrated only multimedia elements for HTML text.

Practical exercises serve to train the interest rules of prefixes use, considered in the theoretical part of the book, and to consolidate the knowledge. Also, the book contains control tests, which can be used to check the quality of the learned material.

Какой префикс подходит к глаголу «бросать»?

1. Он брасывал галстук за спину, чтобы не мешал, а тот всё время падал.
2. Ты хочешь, чтобы я этот чёртов компьютер в окно бросила?
3. Сначала Эмре прошёл половину поля с мячом и аккуратно, нежно бросил мяч через Шеруши, попал akurat в левятку.
4. Они брасывали на деревянную стену большие нечёрные тени и тихо переговаривались друг с другом.
5. То ли я не попал в него, то ли бросил гранату и она не взорвалась,—сказать не могу.
6. За окном золотились деревья во дворе, и ветер горящими брасывал листья.
7. Праздники закончатся, вы вернётесь в обычный ритм жизни и бросите пару лишних килограммов, что набрали за эти дни.

Fig. 2. Exercise to fill the gaps.

In the process of the exercises creation there was revealed imperfection of eXeLearning operation, related to the choice of the size of the text fields to fill the gaps. The program creates the field with the length equal to the number of letters in the missing fragment, which gave a significant hint to a trainee. For solving this problem there was improved the function to create exercises with gaps (generation of an embedded element – instructional device CaseStudyDevice) so that the amount of gaps is not dependent on the missing prefix length.

3 The Tool Tip

Additional element that improves examples understanding used in the tutorial, is the tool tip containing the translation and grammar information of each of the separated words. This data is generated by a server-side programs and transferred to the user's browser, where the only occurs rendering a pop-up window with the information received.

Примеры употреблений

1. Производители косметики высоко ценят растительные масла и вводят их во все кремы: питательные, увлажняющие, регенерирующие, ялифтинг- средства. [Чудо- эликсиры из зернышек берегут твою молодость // « Даша», 2004]
2. Мой **приятель** внёс своё имя в лотерею - проводил спортивный турнир и я выиграл. [Джим Керри -- изнутри (2004) // « Экран и сцена», 2004.03.08]
3. В каждый раз добавляли 50 мкл физиологического раствора, 25 мкл разведённый комплементсодержащей сыворотки, добавляли 50 мкл гемолитической системы. [Н. М. Зубавичене, А. А. Чепурнов. Динамика гемолитической активности комплемента при экспериментальной инфекции Эбола (2004) // « Вопросы вирусологии», 2004.03.08]

Посмотреть динамические примеры

- за-
- за-ся-
- из- (ис-)
- на-
- на-ивать (-ивать)
- на-ся-
- над- (надо-)
- недо-
- о- (об-, обо-)
- о- (об-, обо-) -ся
- обез-(обес-)
- от-
- пере-
- пере-ся
- по-
- по-ивать (-ивать)
- под- (подо-)
- при-
- про-
- раз- (разо-, рас-)
- раз- (разо-, рас-) -ся

Fig. 3. Tool tip in example for the usage of verbs with the prefixes.

A program running in a user browser has access to every single word as to the DOM (Data Object Model) element of the HTML pages on the selector ".tooltip". For the implementation of the tips in the page needs to make a link to a JavaScript file and every word placed inside the tag with the class "tooltip". For this, there was developed a special program «Tagger» in the language of Java, which marks the HTML text and adds the specified tag to each word.

This method allows you to change the contents of the pop-up window and logic of the program, located on the server, without any changes in the content of the tutorial. To activate the tool tip, there was selected a single-click of the left mouse button on the word GUI event. This specific action in a user's browser, from the point of view of the JavaScript program, generates request to the server. For its development there was chosen open JavaScript library jQuery. The library is useful for interaction with the DOM - page elements and construction of HTTP (Hyper Text Transfer Protocol) requests. The creation of HTTP requests to the server is carried out using the standard function ajax(), from jQuery library, working on the basis of the Ajax technology. To synchronize queries the corresponding option has been specified «async: false». This request executes server application which is handling the data.

As the source of the translation there was chosen service MyMemory [5] with an open API (Application Program Interface). Tutorial with using the JavaScript program is able to do the HTTP request and get the response in the JSON format (JavaScript Object Notation), which is easily handled by standard JavaScript methods.

Correct grammar information creates on the server side using Russian language stemmer for non-commercial use - Mystem [6]. This is a separate module that runs as a console application. Commercial program Mystem is not spread and is not planned to. The obtained morphological information is processed by a special server application, written in C language. It provides communication with the server on the

CGI standard (Common Gateway Interface), as well as to retrieve the necessary data and converting them into JSON format.

On the basis of the architecture of the developed server application there is a separate system thread, which is activating a console program Mystem. Output of the program is recorded and processed. The information obtained is divided by basic form of a word or a lexeme and grammar information to this the lexeme. This information is structured for its further use in accordance with the JSON format of data exchange.

As a result there was developed a functioning tool tip that appears when you make a single-click by the left mouse button on the selected word and containing the translation of the word in the English language and its grammar analysis.

4 Actual Examples from the Russian National Corpus

Main content of the tutorial included materials selected by the originator for illustration the usage of words. Additionally, we organized the ability to get the «live» examples from Russian National Corpus by clicking the button. Such «live» examples are extracted in the real-time mode from renewing corpus, and therefore the content always varied, timely and actual examples of words use.

The screenshot displays a web interface with a list of four Russian sentences. A tooltip is visible over the word 'проводить' in the first sentence, showing its grammatical forms: 'V-прош,ед,изъяв,жен,сов'. To the right, a sidebar lists various prefixes and their corresponding forms: 'по- (-ивать) (-ивать)', 'под- (подо-)', 'при-', 'про-', 'раз- (разо-,рас-)', 'раз- (разо-,рас-) -ся', 'с- (со-)', 'с- (со-) -ся', and 'у-'. Below this list, there are checkboxes for 'Употребление префиксальны...' and 'Контрольные тесты'. At the bottom of the sidebar, a 'Lesson progress' section shows 'Progress: 20.31%' and 'Completion rules completed: 0 out of 1'.

Fig. 4. Actual examples from Russian National Corpus.

The program, which extracts the actual examples from Russian National Corpus for the electronic tutorial of the Russian language, should work on the server and comply with a CGI standard. The task of creating such a server application, with the help of the Python language and not using external tools, was divided into separate sub-tasks.

The first sub-task - getting search results from the Russian National Corpus, solved by sending the specified HTTP request to the server of the Russian language and parsing coming response from it. We illustrate this sub-task with the request, which extracts examples of verbs use with the prefix «в» and representing the orientation of the action inside «вбить, втолкнуть, внести, воткнуть, всунуть» (hammer, push, bring in, poke in, stick into). Such a word is used to control the accusative or denoting

a case of nouns with the preposition «В» (into) – «В комнату» (into the room), and we need to find this phrases in the corpus. As an example of the word usage we extract the whole sentence from the corpus.

To prevent finding participles and verbs in the subjunctive and imperative mood, the lexico-grammatical search provides a limitation on the grammatical features: the verb in the indicative mood. We specify the step search of a preposition from the verb (one or two words), because after the verb and before the preposition we often use a direct object (pushed HIM/BOY into the room).

Each of the section in the tutorial has its own query to RNC. The query was developed and optimized, thus, to always receive a request, the most relevant to a specified section in the tutorial. It is possible mistakenly to include in a tutorial small percentages of examples. This is a downside to use «live» search examples from Russian National Corpus. This shortcoming was minimized with the help of a multiple-page viewing search results in repeated request to the server.

At the second stage, we selected interesting examples from server results. For this, we use the settings: select a single example of one of the author and ordered, for example, by date of creation. Also we set the number of examples on the page and requested variant with accents. The latter doesn't always work, if taken examples are not from the main corpus, but from a newspaper sub-corpus, accents are usually absent. The most suitable examples are in the training sub-corpus. It is the most fully marked up, but its capacity is very limited and not always sought examples are present in sufficient quantity.

Response from Russian National Corpus is stored and transmitted to the input of a finite state machine, which implemented the second task – the obtained results analysis to create actual examples. State machine is designed to extract examples and highlight the key words and cutting off unnecessary information. For the implementation of each of functional capabilities, we use a separate pair of states of a finite state machine. Alphabet of a finite state machine is HTML tags and character set of any human language in the used encoding.

A third sub-task - output of search results structured as a HTML text. It was implemented in such a way that finite state machine extract the various types of words. For each of the type of word we have a method of realization in HTML. These chunks are joined and added to the template of the page. After it we create the page and transfer information to the client computer.

5 Integration to LMS

For organization of web access to this tutorial it was integrated in the LMS (Learning Management System). We chose eFront LMS, because its core distributed as an open-source project. Also the university of High School of Economics chose eFront as its system for electronic distance learning. The eFront supports import learning courses as SCORM content.

The program eXeLearning allows any academic specialists creating web content without the need to become proficient in HTML or XML markup. It is a desktop

program for developing resources which can be exported in IMS Content Package, SCORM 1.2, or IMS Common Cartridge formats or as simple self-contained web pages. After filling the tutorial with information using eXeLearnig editor we choose option File - Export and SCORM 1.2 specification. It creates a zip package contains linked HTML files with JavaScript files and CSS styles and its metadata for course structure. The import function to the simple self-contained web pages is used by us for standalone version of the tutorial. The tool tip and actual examples are available if the Internet is connected.

Special feature, like a tool tip required the special HTML structure because a user browser has access to every single word as to the DOM element of the HTML pages on the selector ".tooltip". We add this special structure during filling information using integrated HTML editor and special program «Tagger».

Content or features developing sometimes needs to change page structure. It is impossible to change it partly. Then we need to modify the content of the tutorial we should modify it in the eXeLearnig and reimport it into the Learning Management System. It is inconvenient and single-user procedure and we need to organize a store for latest eXeLearning project.

With the use of static HTML files all functions work directly in a web browser. This imposes some functionality restrictions of the tool. Then we use the toll tip in the standalone version, program get cross domain request to the server for the toll tip information. Not all the security policy of the web browsers allows doing a cross domain request to get information.

6 Future work

After some period of time we realized that HTML tutorial structure limits our capacity for features development. Other way to integrate tutorial to the Learning Management System is to save information to the LMS database as a basic content. This way allows us to process data online in a more flexible way.

We have a plan to import information into the database and organize features with the help of server scripts not after GUI event. We need to do some changes to the developed server applications to renew online features in the tutorial. This will take a few steps.

First, server application will get a text from database, extract words from text and wrap it into the tag with some class. Now it happens when we input information to the eXeLearning with using desktop application «Tagger». Second, application will morphologically analyze this lexeme by stemmer based on Mystem module and translate it with using MyMemory web service. Generated information will include English translation, lexeme base form and morphological data. This information will be added to the corresponding tag as an attribute and the text with new HTML markup will save to the database.

Described architecture allows us to generate standalone version of the tutorial which can support helpful toll tip without Internet connection. For this we need to extract saved information to the tutorial pages and save all pages to the package with

the structure needed and CSS styles. All information will be stored in HTML markup. JavaScript file will just get the data from tag attribute and render it to the tool tip.

This summer the Yandex Company has released a new version of the search application in the Russian National Corpus. Now, it is not only the primary user interface for viewing search results, but a programming interface is available. Programming interface generates same search results as a user interface to which is added additional information. Each of word is accompanied by grammar information structured as a XML file. It contains lemma, grammar analysis, semantic analysis and derivation. New release allows us not to use stemmer Mystem to get grammar information for texts from RNC. But this stemmer helps us to parse texts from other sources.

7 Conclusion

No doubt, the theme of the development of electronic tutorials is now urgent. With the advent of distance education in the electronic form, its active implementation in higher education institutions and in the business sector, we need new quick and cheap approaches to the knowledge representation. In addition, there was formed the demand for high-quality linguistic applications, which would have all necessary for this type of product functional capabilities.

This developed technology can be useful for electronic tutorial creation in any foreign language. It uses a free XHTML editor for the formation of the tutorial content and a set of server applications for realization of the necessary functionality. Server applications have been written in the language Python and C without use of external libraries. A program running in the user's web-browser developed in JavaScript using the open library jQuery, and interacts with the server applications. The use of linguistic corpus, rich in all kinds of different texts, helps to make the study of linguistic material and use of the words more clear for the student.

8 References

1. The Federal Law, February 28, 2012. N 11-FZ «On amendments to the Law of the Russian Federation «On Education" in the part of application of electronic learning, distance learning technologies», <http://kremlin.ru/acts/14625>
2. The eXe project: eXeLearning, <http://exelearning.org/>
3. What is corpus? Site of Russian National Corpus, <http://www.ruscorpora.ru/corpora-intro.html>
4. How to put the accent? The wordmania. Words about words, <http://www.slovomania.ru/dnevnik/2007/08/11/how-to-use-stress-sign/>
5. API technical specifications. MyMemory, <http://mymemory.translated.net/doc/spec.php>
6. Technology. About program Mystem. Yandex company, <http://company.yandex.ru/technologies/mystem/>