"THE VISUALIZATION OF A CORPUS-BASED STUDY OF THE TEXT OF THE MEDIEVAL CHRONICLE"

Соломянник Е.А. (Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург)

Abstract. This study attempts to visualize and popularize the analysis of the text of one of the most important sources of ancient Russian history – the Laurentian Chronicle – on the subject of linguistic heterogeneity. The possible variants of visualization are shown, and the choice is made, taking into account both the usefulness of this visualization in narrow circles of researchers of Ancient Russian literature, and the possibility of its use by a wide range of people. Possible ways of development and expansion of the project are presented.

Intriduction. Russian chronicles are one of the main sources of ancient Russian history, they are certainly the main reference point in the construction of the historical narrative, determining dates, features of relationships and family ties.

It is quite obvious that the chronicles were primarily a kind of manual for the princes in their political and military affairs, and that the mood of the chronicler strongly influenced the way and with what evaluative coloring the events that took place would be described.

It is important to understand that the chronicle is a kind of "collective" work, on which many chroniclers worked in turn, replacing each other. Events in the chronicle text are very often not arranged in chronological order, there may be instructive "inserts" of events from the past. Moreover, since the chroniclers differed from each other in their theological views and political sympathies, the same actions of the prince, for example, can receive a positive assessment in one article, and a negative one in another. Therefore, it is important to divide the text of the chronicle into parts that belong to the pen of different authors, and to understand not only the events described in it, but also the context that existed around the chroniclers and influenced the content.

One of the most promising methods of studying chronicle texts is the method of linguistic stratification (heterogeneity). The essence of this method is to select a group of grammatical features present in the text according to certain criteria, and on the basis of their uneven distribution in the text, place chronological "seams" in it and give them a historical explanation. The method of linguistic stratification is primarily good because there is no need to use comparative material (different copies, editions of the text), since it examines grammatical features in the text.

Corpus linguistics is currently one of the most popular branches of linguistics. The corpus consists of texts selected according to certain criteria and collected together, which are later used in linguistic research. Within the National Corpus of the Russian Language, there is an Old Russian sub-corpus, which contains the linguistically marked text of the Laurentian Chronicle.

By means of linguistic stratification, the Novgorod, Galician-Volyn and Kiev chronicle traditions are already being studied. Until now, the Vladimir-Suzdal chronicle, presented in one of the most famous sources – the Laurentian Chronicle, remains practically unexplored.

In order for the results of the study to find their application not only in narrow circles of researchers of ancient Russian literature, but also among the general public, it is necessary to visualize them.

Research purpose. Initially the idea was to visualize a study in the format of the published text of the chronicle with different grammar characteristics highlighted with different colors (for example, a noun – red color, a verb in the aorist – blue color, a verb in the imperfect – green, etc.). However, such image would not make much sense, as the distribution of grammar features is illustrative only in the entire volume of the text and within one page this rendering would not bear any meaning. There would also be certain difficulties about the designation of compound verb forms, for example, plusquamperfect, which can be formed using an auxiliary verb in aorist, imperfect, or perfect form. Starting first and foremost from understanding that it is necessary to somehow represent the entire volume of the text of the chronicle, I was able to choose the most suitable pattern – "timeline" with

the years in chronological framework of the study period, that is, from the XIIth to early XIIIth centuries. At the same time, it was very important to make a visualization interactive, so that a user could hover the cursor over the mark on the timeline and see what kind of grammatical feature was presented, in what context it was used, the full grammar description and position in the text (a column and a row).

So, it was decided to use the Timeline.JS service developed by Northwestern University (USA), which provides all features of the above.

Results. As mentioned above, first of all, such visualization will help to expand the circle of people who will be able to work with the text of the chronicle, since it makes it optional to have an absolute knowledge of the grammar of the Old Russian language to understand the concept of the chronicle. In addition, such visualization can be a useful visual aid for students studying Old Russian and Old Slavonic languages at the university: for example, the construction "dativus absolutus" often causes difficulties, and the experience that can be acquired by using such a manual will certainly help to learn how to recognize this construction in the text and to translate it into modern Russian, depending on the context, properly.

Thus, the ideal option for implementing the research results in the user environment is to create a special website dedicated to the Laurentian Chronicle (in the future, the website can be expanded, and other sources of ancient Russian history can be added there), with the described visualization published on it, which can be developed by adding a search for grammar features, and other features, in addition to those that were directly used in this study.

Соломянник Е.А. (автор)   Подпись