The Informational and Semiotic Nature of French Socio-Political Terminology and its Thesaurus Modeling*

Yurij Gorbunov
yourigorbounov@tltsu.ru

Evgenij Gorbunov
evgenigorbunov@yandex.ru

Sergei Mkrtichev
sm5006@yandex.ru

Togliatti State University, Togliatti, Russian Federation

Abstract

This article is devoted to the analysis of modern ontologies and thesauruses used in the field of professional education. The main attention is paid to the description of the informational and semiotic nature of French sociopolitical terminology, the development of linguistic technology for compiling a terminological thesaurus, reflecting the semantic content of the subject area of French sociopolitical translation, and the study of the semantic relations of French sociopolitical terms. The practical value of this study is to develop a teaching French-Russian thesaurus of social and political terminology, which is proposed to be used in the educational process as a computer linguodidactic tool for the formation of professional translation competence of linguists.

Keywords: ontologies, thesauruses, thesaurus modeling, informational-semiotic nature of terminology, socio-political terminology, automatic translation, professional translation competence of linguists

1 Introduction

This article discusses the preliminary results of the implementation of the project, which provides for the study of the informational-semiotic nature of the French sociopolitical terminology and its modeling in the form of a teaching thesaurus. In accordance with the project, the study is carried out in two stages: at the first stage, the “Teaching French-Russian thesaurus of sociopolitical terminology” is modeled, at the second stage it is tested.
and introduced into the educational process as a computer linguistic and didactic tool for the formation of professional translation competence among students studying in the direction of preparation of the bachelor “Linguistics”, profile “Translation and Translation Studies”, in the subject area of socio-political translation.

The rapid development of linguistics during the twentieth century, many areas, schools, concepts that have developed within its borders, contributed to the active process of formation of thesaurus lexicography. A new turn in the development of thesaurus lexicography occurred in connection with the active development of computer linguistics, when the practice of translation included industry automated and automatic dictionaries. There are new ontologies that contain special knowledge bases that reflect the conceptual content of certain subject areas [Beliaeva 2016, p. 59].

Ontologies form a kind of Foundation that underlies the linguistic construction of thesaurus for individual branches of knowledge. Special attention is paid to the problems of thesaurus modeling of industry terminology using new computer technologies in the field of education.

2 Ontologies and thesauruses in the educational sphere

New ontologies and thesauruses find practical application in the educational sphere. In particular, we consider the possibilities of algorithmization of ontology development based on intermediate models-intelligence maps and conceptual maps [Artemova et al. 2015]. For automated optimization of conceptual maps, it is proposed to use metrics, including subjective metrics of graph topology research. Based on the obtained calculations, recommendations for automated optimization of the considered conceptual maps are proposed, namely:

- remove the vertices of the concept map if their number is excessively increased, making it easier to perceive;
- remove / add vertices of the concept map in the form of a tree, which contributes to a better balance of the concept map;
- remove / add connections on the conceptual map not in the form of a tree, which significantly affects the connectivity of graph simplexes in the complex [Artemova et al. 2015].

Thus, this study justifies the use of metrics for automated optimization of conceptual maps in the development of ontologies in the field of education, and also indicates the place of the stage of selection and optimization in the General algorithm of ontology development based on these intermediate models.

Since the volume of information in the field of education is constantly growing and new information technologies are developing, this entails the search for new ways to optimize these processes. One possible solution to this problem is a thesaurus approach. The concept of thesaurus is widely used in various fields of science, including artificial intelligence and computer science. However, the term "thesaurus" has an interdisciplinary character, so it is found in various fields of knowledge such as logic, philosophy, psychology and pedagogy, linguistics.

The main advantage of the thesaurus approach in relation to others is that it allows you to “sort out” a large amount of information for better orientation in it and form a subject thesaurus in a specific subject area, adding to it and managing it [Tulokhonova 2017]. Modeling of the learning system based on the thesaurus approach involves the formation and development of the thesaurus of the student in a specific subject area based on managed
dictionaries through the use of taxonomy. In this case, the thesaurus is considered as a system of knowledge of a certain subject area (discipline), structured in a certain way and presented in the form of concepts and semantic relations between them.

Thesaurus as a special form of representation of knowledge of a particular subject area is easily amenable to automated processing, which ensures the optimization of the educational process. However, in this case, there is a problem of choosing the most effective from the point of view of the educational process automatic methods. Thus, in the course of research on automated construction of the lexical core of the thesaurus in a narrow subject area, this problem is solved by using distributive and statistical analysis of a large body of texts (about 15 billion word usage) using the tools of the Sketch Engine system. For paradigmatic and syntagmatic relations of lexical units thesaurus calculates a quantitative value of the coupling strength and study the effect of enclosure volume on the quality of the thesaurus [Zakharov 2015].

In this study, it was shown that modern corpus linguistics has an Arsenal of tools that allow automatically based on the corpus of texts to form classes of lexical units that are in relation to ontological dependence. In particular, the example of the term "engine" revealed a set of lexical units that form the nest of the thesaurus, or lexico-semantic field around this term. In this case, the use of the corpus of texts and tools of the Sketch Engine system allows us to identify syntagmatic and paradigmatic connections in an automated mode and provide adequate content of the lexical and semantic system. In this study, was demonstrated the relationship between trunk size and quality (accuracy) of the generated thesaurus, namely the larger the enclosure volume, the higher the quality (accuracy) in a given volume of distribution thesaurus. Since concepts in the conceptual scheme are expressed not only in individual words, a mechanism was used to identify stable phrases (collocations and colligations), many of which are terms, and they were included in the thesaurus. The use of different approaches to the formation of the thesaurus allowed to achieve a high degree of lexical saturation. It is obvious that the resulting thesaurus is an input for the expert, whose task is further to describe the relationship between these words with the title word and each other.

In the future, it is planned to compare thesauruses for the same subject area, formed for the same keywords, but on different types of cases: on the one hand, on a large General language case, on the other — on a specialized case. It is assumed that the latter will give the best result. But the former have the advantage of a large volume. To implement such comparisons, it is necessary to use metrics that can be used to obtain a numerical value of the accuracy and completeness of the thesaurus. In addition, it is necessary to improve the software and linguistic support of the buildings. According to the researcher, it is also necessary to improve the quality of morphological markup, since lemmatization errors lead to the fact that word forms of the same lexeme are considered autonomously, which distorts the statistics. Among other things, the grammar of lexical and syntactic templates, which is based on distributive and statistical analysis in the Sketch Engine system, needs to be improved [Zakharov 2015].

After a brief review of individual technologies and approaches used in the educational sphere, we will proceed to the analysis of preliminary results of the study of the information-semiotic nature of French socio-political terminology using thesaurus modeling.
3 Thesaurus modeling of French socio-political terminology

The study of French socio-political terminology has deep roots and long linguistic traditions. In this direction, a great interest is the thesaurus approach to modeling socio-political terminology, which allows not only to study the information-semiotic nature, but also to systematize and classify socio-political terminology for the purpose of subsequent use of the results in the educational process of forming professional translation competence of students-future bachelors of linguistics. This thesaurus is based on lexicographic traditions that have developed in the international scientific group “Speech Statistics”, the founder of which is doctor of Philology, Professor R. G. Piotrovsky (RSPU them. A. I. Herzen, St. Petersburg) [Piotrovsky 1984; 1999; 2006]. This also takes into account significant experience in the development of the General theory of linguistic dictionaries-thesaurus [Karaulov 1981; 2010], used the achievements of applied linguistics in science and education [Beliaeva 2007; 2010; 2016], new information technologies in linguistics [Zubov, Zubova 2004], as well as the results of research in the field of computer linguodidactics [Piotrowska 2005].

The main goal of our study is to identify and describe the informational-semiotic nature of the French socio-political terminology using thesaurus modeling. The main objectives are following:

1) to carry out linguistic construction of the main semantic areas of French socio-political terminology;

2) to identify the semantic relations that form the structural network of the Teaching French-Russian thesaurus of social and political terminology;

3) to reduce individual semantic areas and fields into a single semantic space of the Teaching French-Russian thesaurus of social and political terminology.

The scientific novelty of this study is that for the first time it is proposed:

1) to investigate the informational-semiotic nature of French socio-political terminology using thesaurus modeling;

2) to develop principles for modeling French socio-political terminology in the form of a thesaurus with a hierarchical structure;

3) to build the Teaching French-Russian thesaurus of socio-political terminology, which is a complex informational-semiotic model that includes the French terminology and terminology system with an “exit” into the Russian language;

4) to reveal the semantic relations of French socio-political terms that connect them into a single semiotic system and make up its structure;

5) to develop the information base of the teaching linguistic automaton in French special terminology, which can be the basis for the multidimensional teaching model “Thesaurus of the linguist-translator” with the aim of forming professional translation competence among students - future bachelors of Linguistics.
The goal and objectives of the study are realized with the usage of a comprehensive methodology, which includes analysis of modern lexicographic sources in French socio-political terminology, a method of componential analysis of definitions of French socio-political terms, and the method of linguistic construction of a thesaurus model of the subject area being under study.

At the first stage of the study, a componential analysis of the definitions of French socio-political terms is carried out; semantic relations between the terms are revealed. The results of a componential analysis of definitions of French socio-political terms are systematized and classified in a thesaurus model. A componential analysis of definitions of socio-political terms is carried out according to explanatory dictionaries, as well as done on the basis of author’s definitions of socio-political terms according to specific scientific sources, such as: scientific articles, monographs, periodicals in French. A componential analysis of French socio-political terms allows us to establish the term field and the term system of the studied subject area. As follows from the synoptic scheme, the Teaching French-Russian thesaurus of socio-political terminology includes 5 large semantic areas, namely:

- Semantic area 1: “Société et organization politique” - “Society and political organization”;
- Semantic area 2: “Le droit” - “Law”;
- Semantic area 4: “La communication et l’information” - “Communication and information”;
- Semantic area 5: “La guerre et la paix” - “War and peace”.

Each semantic area is divided into several semantic fields, which combine in their space terms and terminological phrases based on a common integrative semantic feature.


Each semantic field of this thesaurus is a separate vocabulary article in which socio-political terminology is based on a common semantic feature, taking into account the semantic relationships identified in contextual uses. The dictionary entry for a separate semantic field is constructed according to the following scheme:

- the title term, reflecting one of the key concepts of the term field of the subject area and its Russian equivalent;
- variants of the term in French and their Russian correspondences;
- synonyms of the French term and their Russian equivalents;
- antonyms of the French term and their Russian correspondences;
- a hyperonym of the French term with translation into Russian;
- hyponyms of the title term and their Russian equivalents, classified by semantic, formal and syntactic criteria;
- meronyms of the French term and their Russian equivalents;
- an indication of the semantic relations specific to the title term with other terms of the given branch of knowledge, such as the relationship of “alternative”, “belonging to theory”, “correlation”, etc.

In the process of constructing the Teaching French-Russian thesaurus of socio-political terminology, the linguistic nature of the semantic relations of terms is studied, among which a special place is occupied by relations of variance, synonymy, equivalence, antonymy, hyper-hyponymy, meronymy, correlation, and others. Semantic relations combine the studied terms into a single term system and form the semantic structure of the Teaching French-Russian thesaurus of social and political terminology. Let us consider in more detail the characteristics of the identified semantic relations of French socio-political terms.

1. **Hyperonymy (hyperonymie)**: from hyperonyme - hyperonym: a concept in relation to another concept, expressing a more general essence. So in the subject area under consideration, the term “la politique” (“politics”) is a generic term, or a hyperonym in relation to the terms “la politique intérieure” (“domestic policy”), “la politique extérieure” (“foreign policy”), “la politique colonialiste” (“colonial politics”), “la politique conservatrice” (“conservative politics”), “la politique expansionniste” (“expansionist politics”), “la politique libérale” (“liberal politics”). The semantic attitude of hyperonymy is also found in the semantic field of the French term “visite” - “visit”, which acts as a hyperonym in relation to its hyponyms: “la visite d’Etat” (“state visit”), “la visite de travail” (“Working visit”), “la visite officielle” (“official visit”), “la visite privée” (“private visit”).

2. **Hyponymy (hyponymie)**: from hyponyme - hyponym: a concept expressing a particular entity in relation to another, more general concept. So the terms “la guerre civile” (“civil war”), “la guerre d’agression” (“predatory war”), “la guerre de défense” (“defensive war”), “la guerre froide” (“cold war”), “La guerre fratricide” (“fratricidal war”) are related by the relation...
of hyponymy to the term “guerre” (“war”), which is their hyperonym. The semantic attitude of hyponymy is also observed within the semantic field of the French term “police” (“police”), to which its hyponyms obey: “la police judiciaire” (“judicial police”), “la police militaire” (“military police”), “la police politique” (“political investigation”), “la police privée” (“private investigation”), “la police secrète” (“secret police”), “la police de l’air” (“air police”), “la police des moeurs” (“morale police”), “la police de la route” (“traffic police”).

3. Meronomy (méronymie): this is a classification of phenomena based on the relationship of meronomy and holonomy. Meronymy and holonomy as semantic relations are mutually connected to each other. A meronym expresses a concept that is an integral part of another, and a holonym expresses a concept that is a whole over other concepts, so another concept is part of the first. For example: the term “corps de police” is a holonym in relation to the terms: “agent de police” (“policeman”), “commissaire de police” (“police commissar”), “ministre de la police” (“Minister for Police Affairs”), “préfet de la police” (“prefect (chief) of the police”), “indicateur de la police” (“police whistleblower”), which are his meonyms.

4. Synonymy (synonymie): synonyms – are words of one part of speech, different in sound and spelling. For example, the term “les pays émergents” (“developing countries”) acts as a synonym for phrases such as “les économies émergentes” or “les pays d’émergence économique” (“emerging economics”). The semantic relation of synonymy also connects the terms “négociations”, “pourparlers” and “entretiens”, which are translated into Russian unambiguously: “negotiations”. Examples include: “mener des négociations” - “to negotiate”, “entamer les pourparlers” - “to start negotiations”, “prolonger les entretiens” - “to extend negotiations”.

5. Antonymy (antonymie): a relationship found between terms that are opposite in meaning, for example: “la paix” (“peace”) - “la guerre” (“war”). The semantic attitude of antonymy is also observed in contrasting terms “la convergence des points de vue” (“coincidence of points of view”) and “la divergence des points de vue” (“divergence of points of view”).

6. Correlation (corrélation): from lat. correlatio - correlation, interconnection. So, for example, the term “la classe ouvrière” (“working class”) is connected by the semantic correlation relation with the terms “la lutte de classe” (“class struggle”) and “les rapports de classe” (“class relations”). In the same way, there is a correlation in the chain of terms “le prolétariat” (“proletariat”) and “la dictature du prolétariat” (“dictatorship of the proletariat”).

A distinctive feature of the dictionary article is that it possesses an open character, that is, taking into account special goals, the dictionary article can be expanded with some new terms using special notes that introduce a new thesaurus terminology reflecting current trends in the evolution of the subject area. In the process of independent work on socio-political translation using this vocabulary-thesaurus, the student must follow the translation algorithm, which includes the following operations:

1. Searching for the French socio-political term in the thesaurus index.
2. Determining the addresses of semantic fields, which include the French socio-political term.
3. Consistent access to the semantic fields of the French term at the specified addresses.
4. The choice of the Russian equivalent of the French term as a result of the analysis of its semantic fields, taking into account the real context.

As a result of repeated reference to the various semantic fields of a particular French socio-political term searching for the corresponding Russian equivalent, the student forms a professional thesaurus that reflects the corresponding subject area. Over time, in conditions of practical training in socio-political translation, the French-Russian thesaurus of socio-political
terminology is fixed in the professional consciousness of the future translator and ultimately becomes an indicator of his professional translation competence as a specialist in the field of socio-political communication.

The Teaching French-Russian thesaurus of socio-political terminology is intended for practical use in the educational process of a university as part of an educational and methodical complex that includes, in addition to the thesaurus itself, a textbook on socio-political translation from French into Russian [Gorbunov, Gorbunova 2016]. This study guide contains relevant training material that reflects the main trends in socio-political life in the world, as well as in France and Russia. Significant attention in this manual is given to the organization of students’ independent work in preparation for translation practical classes for the purpose of deep study and mastering of the terminological material of the studied branch of knowledge based on its thesaurus model.

4 Conclusions

So, the informational and semiotic nature of French socio-political terminology is revealed in the form of a thesaurus model of the semantic space “Society and Politics”. The Teaching French-Russian Thesaurus of Socio-Political Terminology, constructed in current study, is a complex informational-semiotic model that includes the French terminology and terminology with an “exit” into the Russian language. The semantic fields that make up the term field form the basis of the vocabulary article of the training thesaurus. The system of French socio-political terms is consolidated into a single semantic space due to the structural network of semantic relations of hyper-hyponymy, meronymy, synonymy, antonymy and other associative relations.

The practical value of this research is to develop a French-Russian thesaurus of socio-political terminology, which is proposed to be used in the educational process as a computer linguodidactic tool for the formation of professional translation competence. The proposed method using a special terminological vocabulary-thesaurus not only contributes to the effective study of the content of the subject field “Society and Politics”, but also promotes the development of special humanitarian thinking among students, which is necessary solving practical problems in the future professional activity of a linguist-translator in the subject area of special socio-political translation.

The proposed thesaurus can also be used in the development of other specialized online translators and dictionaries. The creation of a “Teaching French-Russian thesaurus of socio-political terminology” and its introduction into the computerized educational process will allow us to make significant progress in the training of qualified linguists and translators.

We are planning to organize educational and practical work on building thesauri in such well-known collaborative development platforms as Protege and Vocbench.

References


