NAVIGATING LANGUAGE: THE THESAURUS AND ITS IMPACT ON LINGUISTIC EXPLORATION

Akbarova Marziyahon Akbarjon qizi

Farg'ona Viloyati , Qoqon shahri, 2 ixtsoslashtirilgan maktabi

Navigation Autonomy of foreign language learners provides a new understanding of the theory and practice of student autonomy in the context of foreign language teaching and does so in multiple languages and with the help of multiple voices. The authors demonstrate effective research methods and new directions, as well as report on the state of student autonomy in institutions around the world. Most authors write about their experience of realizing the autonomy of foreign language learners at home or in the dominant language(s). The volume contains complete chapters or excerpts in 15 languages: Czech, Danish, English, Finnish, German, Hungarian, Italian, Japanese, Korean, Mandarin, M. D., Portuguese, Spanish, Thai and Turkish. Each chapter is accompanied by a chapter or summary in English, a glossary and a few thoughtful questions. As a starting point, David Little offers a theoretical introduction, and in conclusion, the Editors analyze the participants' stories and interpret the process of managing autonomy using various languages.

Any subject of educational and information activity can be characterized by a hierarchical dictionary of concepts in this field -a thesaurus. In addition, such a thesaurus has long been compiled on many topics. A thesaurus is a tree of concepts on a specific topic, ending at the top, the most general, and at the bottom, the most specific, narrower concepts. The words (terms) in the thesaurus are usually related by the relation general-private, whole-private, etc. In an expanded sense, the thesaurus is interpreted as a description of a system of knowledge about reality that has a separate information carrier or a group of carriers. This medium can act as a recipient of additional information, as a result of which its thesaurus is changed, and the original thesaurus determines the recipient's capabilities when receiving semantic information.

Thesaurus-this is a term widely used in computer science as an integral part of information retrieval systems.   An information retrieval thesaurus is a dictionary of terms and phrases compiled according to certain rules for a specific subject area and designed to improve the quality of information search in a given subject area.

"When they say thesaurus, nowadays they often mean Thesaurus linguae latinae, an enterprise of five German academies started in 1900. Characteristic feature: It includes all the important words that occur in a particular language, and under each word there are examples from texts available in that language.

A.N. Baranov and D.O. Dobrovolsky, in the preface to the book " Horse Edition ", gave the following definition of a thesaurus – "a special type of dictionary that differs from others (in particular, explanatory, bilingual, etc.) in the way of organizing linguistic material. In the thesaurus, the language units are not listed in alphabetical order, as in a regular dictionary, but are grouped by their meaning." In the field of information retrieval thesauri,

there are two main directions: indexing documents using certain rules that use the semantic meaning of the text, and the use of hierarchical, associative and synonymous relationships when processing user search queries. In addition, semantic relationships between descriptors can be used to classify and categorize documents, compile a list of words associated with a query, and some other information retrieval tasks. Its set of terms should include all terminological expressions used in the text. For the practical implementation of this method, the presence of most terms in the thesaurus is sufficient. The description of the natural language of any subject area is characterized by some constructions used for text convolution, such as the same terms, anaphoric and elliptical convolutions. Homonymy creates an additional problem. The same words are associated with anaphoric and elliptical convolutions. Some tasks can be solved very effectively by including an extensive database of synonyms in the source thesaurus, whereas tasks related to solving homonymy problems can be effectively solved only with the help of statistical analysis.

A thesaurus is a standardized vocabulary consisting of basic concepts, a set of semantic relationships defined between them.

Lukashevich N.V. in his book "thesaurus in the tasks of information retrieval", he put forward the following ideas: "creating a thesaurus is not an easy task. Although there are certain standards for creating thesauri, it is not always possible to directly reproduce existing methods. The reasons for this are, firstly, the lack of sources of lexical information (for example, a corpus of fixed texts), and secondly, the specifics of the subtitle of the subject area.

The first problem can be solved by creating a lexicographic resource similar to your own. Corpus linguistics tools are used to solve this problem.

The second task is impossible not to solve, for this it is necessary to design the structure of the thesaurus in such a way that it is easy to adapt the thesaurus to any field. It follows from this that the software tool can hypothesize about the alleged relationships between terms containing errors only with a certain degree of probability as a result of analyzing the dictionary text. Attempts to improve the quality of the program inevitably lead to the fact that the cost of improving it exceeds the cost of correcting such errors manually. The tool implemented to solve this problem should select terms at the first stage of text analysis, and then look for relationships between them based on the analysis of the selected terms. The set of conditions and relationships found at each stage of the work must be verified by the user, who modifies it according to his needs. The set obtained at the final stage of the work is used to create a thesaurus. At work, it is necessary to study methods and tools that will help automate the compilation of a thesaurus. It is appropriate to study the technologies and determine the main stages of the thesaurus compilation process manually, and then suggest ways to automate the process. The development of an automated approach to the thesaurus of subject areas is the main part of the work and is based on a combination of methods for analyzing the corpus of texts of the subject area and manual labor of a specialist.

REFERENCES:

1. (Aktaev, E. K. Oʻquv va axborot jarayoniga avtomatlashtirilgan tezaurusning zamonaviy usullarini joriy etish / E. K. Aktaev.- - Mat: bevosita // Eat olim- 2015. - No. 6,2 (86,2). - B. 3-6. — URL: https://moluch.ru/archive/86/16526/)

2. Lukashevich N.V., Saliy A.D., Demonstration of knowledge in the automatic text processing system // NTI, 2-ser. 1997. No. 3. pp. 1-6.

3. Zhuravlev S.V., Yudina T.N., Information system RUSSIA // NTI, Ser. 1995. No.3. Pp. 18-20.

4. Winston M., Chaffin R., Herman D. The taxonomy of relationships in general // Cognitive Science. 1987. No. 11. pp. 417-444.