

## Theoretical Basis of Thesaurus Dictionaries

BABAZARIPOVA MALAKHAT ERGASH KIZI  
*Navoi State University, Uzbekistan*

### ABSTRACT

*This article analyzes the theoretical foundations of thesaurus and idiographic dictionaries, the principles of their construction and their differences. The importance of dictionaries, their role in scientific and practical fields are discussed. It is shown that idiographic dictionaries are aimed at grouping words semantically and thematically, while thesauruses are aimed at systematizing semantic relations between lexical units, such as synonymy, antonymy, hyponymy, hypernymy. The article highlights the theoretical approaches to these types of dictionaries through the research of such famous linguists as L. V. Shcherba and M. Morkovkin. Also, scientific and practical principles for developing thesaurus dictionaries in the Uzbek language are discussed, their application in the fields of information technology and computer linguistics, and their role in creating an electronic corpus dictionary are emphasized. The differences between thesauri and ideographic dictionaries in terms of their structure, purpose, semantic relationships, and their role in research are presented.*

**Keywords:** Thesaurus, idiographic dictionary, lexicography, semantic relations, thematic grouping, synonym, antonym, hyponym, hyperonym, electronic dictionaries, thesaurus of the Uzbek language, theoretical and practical significance.

### INTRODUCTION

The creation of various types of dictionaries is of universal cultural importance. Dictionaries help each of us to work

correctly and accurately in the scientific process, to think realistically and objectively, and to imagine. Lexicography is a branch of lexicology that deals with the theory and practice of compiling dictionaries. As we mentioned above, there are two main directions in lexicography regarding the location of words:

1. Placement based on pronunciation proximity;
2. Placement based on spiritual proximity

The first method is based on the creation of alphabetical dictionaries. Although these dictionaries organize words, they do not correspond to the systematic knowledge of man about the world.

The second method is based on semantic proximity, and this method is the basis for the construction of thesauruses and ideographic dictionaries, which are based on the principle of combining words that are not at all similar to those in dictionaries arranged alphabetically, according to their semantic proximity. So what are such dictionaries? How are they constructed? A thesaurus is a special dictionary in modern linguistics, or one of the special types of such dictionaries, capable of reflecting semantic relationships between lexical units - concepts such as homonyms, synonyms, antonyms, paronyms, hyponyms, hypernyms. These dictionaries differ significantly from other dictionaries. In particular, they are formed as ideographic dictionaries that reflect the semantic relationships between lexemes within a field of science or units within a subject.

#### RESEARCH OBJECT AND METHODS USED

This article takes thesaurus and ideographic dictionaries as the object of research, and studies their differences, structure, and theoretical foundations. The central object of the study is the creation of these dictionaries in the Uzbek language, their practical and scientific application, in particular, their significance in the field of information technologies and computer linguistics. Methods of analysis, comparison, and classification are used as a methodological approach. With the

help of these methods, the structure of thesaurus and ideographic dictionaries, their role in systematizing semantic connections, and their linguistic features are determined. The historical approach is used to analyze the principles of dictionary creation based on the works of famous scientists. In practical terms, attention is paid to the directions of creating electronic dictionaries and their effective usage.

Ideographic dictionaries are designed to systematize information about lexemes. They differ from thesauri mainly in the principles of thematic or semantic arrangement of lexical units. For example, ideographic dictionaries are aimed at systematically placing terms that cover a certain topic, while thesauri are designed to show the interrelationships of lexical units through semantic connections. Therefore, the question arises: what are the main differences between these dictionaries and what problems can they solve? In answering these questions, the research of famous world lexicologists and lexicographers – L. V. Sherba [2]; H. Kasares [8]; M. Morkovkin [7]; V. V. Vinogradova [9]; O. S. Anatolyevna [6] is of great importance. The work of these scholars play an important role in the theoretical substantiation of ideographic and thesaurus dictionaries and their practical application.

Russian linguistics has a rich experience in the field of ideographic and thesaurus dictionaries. L. V. Shcherba contributes a special place in the development of the theory of these dictionaries. He paid great attention to the organization of dictionaries on a semantic and logical basis. In Shcherba's research, dictionaries are considered not only as a crucial tool for studying language, but also for its systematic analysis [1].

When observing some of the characteristics of thesauruses and ideographic dictionaries, firstly, it is important to see their differences in the following aspects:

1. **In terms of semantic relationships:** Thesauruses are aimed at systematizing semantic relationships of lexical units, such as synonyms, antonyms, hyponyms, and hypernyms. Ideographic dictionaries, on the other hand, provide meaningful grouping within a subject or field.

2. **In terms of structural principles:** Thesauruses can arrange units in a network and show semantic connections between them. Ideographic dictionaries, on the other hand, divide information into thematic sections, ensuring their logical order.
3. **In terms of intended usage:** Thesauruses are often useful in linguistics, artificial intelligence, and information technology, while ideographic dictionaries are more useful for scientific, pedagogical, and cultural purposes.

In particular, L. V. Sherba, in his works, explained ideographic dictionaries and thesauri as special types of dictionaries, emphasizing their features and functions. He indicates that the two types of dictionaries differ from each other significantly in terms of purpose and structure [2].

The *National Encyclopedia of Uzbekistan* says the following about the ideographic dictionary:

IDEOGRAPHIC DICTIONARY – is a type of dictionary. Such dictionaries are not arranged alphabetically, but by subject. [3]

The term thesaurus is defined as follows:

THESAURUS (Greek thesauros – treasure, wealth) – 1. A dictionary that covers all words in a particular language and fully reflects their usage in the text. 2. A dictionary in which lexical units (words) related to a particular field of science are arranged in subject order and the semantic (genus-species, synonymous, etc.) relationships between these units are reflected. In this case, the words are not arranged alphabetically, but all the words related to one subject or concept are given in one place; the required word is searched for by concept. [4]

In A. Hojiyev's *Explanatory Dictionary of Linguistic Terms*, the following explanation is given for the thesaurus:

THESAURUS (Greek: Thesaurus - treasure) A dictionary that fully reflects the lexical content of a language [5].

Also, in Yorkinoy Hamroyeva's Ph.D. dissertation entitled 'Principles of Idiographic Dictionary Creation', information is provided about idiographic and thesaurus dictionaries and their specific features.

#### RESULTS AND THEIR ANALYSIS

Based on the above sources and our personal opinions, we can make the following observations about the specific features of idiographic and thesaurus dictionaries.

##### *Idiographic dictionaries*

1. **Constructed on the basis of concepts:** Idiographic dictionaries group words by semantic proximity or by topic. These dictionaries organize information based on horizontal classification.
2. **Words are searched by meaning:** When using these dictionaries, the user identifies the concept he needs and finds all relevant lexical units within this concept.
3. **Emphasis is placed on classifying rather than linking linguistic units:** Idiographic dictionaries focus on organizing knowledge through hierarchical classification of concepts rather than analyzing the relationships between linguistic units.
4. **They serve mainly for theoretical purposes:** These dictionaries are used more for the study of linguistics and systems of concepts.

##### *Thesaurus dictionaries, on the other hand*

1. **Constructed based on semantic relationships:** Thesauri show synonymy, antonymy, hypernymy, hyponymy, and other semantic connections between words.
2. **It is considered an information retrieval tool:** A thesaurus allows the user to identify semantic relationships between language units. This finds words and their connections related to a given concept.

3. **Mainly used for practical purposes:** These dictionaries are widely used in areas such as translation, information retrieval, and automatic analysis.

We can also clearly see the differences between idiographic dictionaries and thesauruses in Table 1 below.

Features	Idiographic dictionaries	Thesaurus dictionaries
<b>Structural principle</b>	Hierarchical classification of concepts	Demonstrating semantic relationships
<b>Purpose of use</b>	For theoretical analysis and scientific research	For information search, translation and data systematization
<b>Search for information</b>	Search for words based on a specific concept	Identifying relationships between words
<b>Classification method</b>	Horizontal: by concepts	Vertical: by the relationship between words
<b>Field of use</b>	Linguistics, lexicography	Applied linguistics, computational linguistics

Thus, L.V. Shcherba emphasizes that each of these dictionaries is vital in its own right and he says that it is necessary to take a different approach to study lexical units and according to his idea, idiographic dictionaries are intended for more theoretical knowledge, thesauruses are considered a convenient tool for solving practical problems. So, a thesaurus and ideographic dictionaries are important in systematizing language and determining the interrelationships of lexical units. In-depth study and analysis of their specific features further enriches the theory of dictionary construction and increase their practical significance.

Therefore, each of idiographic and thesaurus dictionaries has its own theoretical basis and practical application, and distinguishing between them plays an important role in clarifying the principles of scientific research and dictionary creation . The great linguist scientist V. V. Morkovkin also shared his

experiences with others about the procedure, guidelines and legal provisions for compiling dictionaries with thesauruses in his article 'К вопросу об учебном тематическом словаре' (On the issue of an educational thematic dictionary). In this article, Morkovkin compares thesauruses and thematic dictionaries [7].

In fact, from the outside, these two dictionaries seem to be repeating each other, but this does not mean that they are exactly the same thing. In a thematic dictionary, lexemes are grouped around a specific topic and are united within a specific topic. When developing a thesaurus dictionary, it is necessary to clearly distinguish between words that are united within a specific topic and serve only this scope. This situation causes the separation of concepts such as "topic" and "situation." For example, the word "ice cream" can be considered a member of "the dessert" (pie, cake, ice cream, candy) group. This word can also be seen in the park, alley, kiosk, ice cream, bench ... series. The first series of words is thematic series, and the second is a series of situationally related words. According to Morkovkin, in a thematic dictionary, words, regardless of their category, should be combined into thematic rows.

Based on the above considerations, the difference between thematic and thesaurus dictionaries can be defined as follows:

- Thematic dictionaries are made up of groups within a specific topic, classifying words in the language according to these groups and combining them within these groups.
- Dictionaries with thesauruses group together a certain concept, its synonyms, antonyms, and words related to that lexeme.

Since our research is mainly focused on thesaurus-type dictionaries, we will pay detailed attention to the theoretical foundations of this type of dictionaries and the principles of their organization. In particular, we will study the role of thesaurus in systematizing semantic relations between lexical units, as well as important aspects of their scientific and practical significance.

Thesaurus-type dictionaries are important in creating electronic corpus dictionaries. Using electronic corpus

dictionaries to quickly and easily find information can achieve very good results.

N. V. Lukashovich emphasizes the complexity of the process of creating a thesaurus dictionary. In his work *Тезаурусы в задачах информационного поиска* (Thesauri in information retrieval tasks), he cites theories of thesaurus construction. According to him, although there are separate standards developed to implement this process, applying them in a uniform manner often creates complexity. Such problems are mainly related to two main factors:

1. **Lack of lexical information sources:** for example, insufficient corpuses of texts or other lexicographic resources. To solve this problem, it is necessary to create new lexicographic sources. Corpus linguistics tools can effectively help in this process.
2. **The specific features of the subject area:** Since each field requires a tailored approach, the structure of the thesaurus must be flexible.

These factors can limit the technologies used in thesaurus construction. Software tools can identify approximate connections between terms by analyzing dictionary texts, but this does not exclude the possibility of errors. The cost of improving the software often exceeds the cost of manually correcting these errors [11]. To successfully organize the thesaurus construction process, it is necessary to use the following approaches:

- The initial stage of text analysis involves identifying terms and analyzing them based on semantic relationships;
- The relationships found at each stage are verified by the user and modified to suit their needs;
- Using the information obtained in the final stage to create a complete thesaurus.

It is also important to explore the possibilities of automation in this process. Identifying the main stages of the manual

compilation process and implementing methods for automating them will allow for the efficient creation of thesauri.

The automated approach to the construction of subject-specific thesauri focuses on the use of corpus analysis techniques and the integration of manual work by experts. This approach involves the implementation of the following main tasks:

1. Developing research methods for compiling a thesaurus;
2. Research methods aimed at automating thesaurus construction;
3. Developing the structure of the thesaurus;
4. Developing methods for extracting terms from a text corpus;
5. Developing new methods for identifying semantic relationships;
6. Development of system architecture and functional diagram for building thesauri;
7. Testing creative tools and evaluating their effectiveness.

#### CONCLUSION

Based on the above considerations, we can say that compiling a complete dictionary of the Uzbek language with a thesaurus and developing its electronic versions will take the lexicography of the Uzbek language to another level.

Another important aspect of thesaurus dictionaries is that they can be of great importance in the comparative study of different languages. In addition, thesaurus dictionaries can also be of great importance in the development of translation dictionaries. For example, something that is called by one word in one language can be called by several different words in another language. Consulting a thesaurus is a very interesting activity, especially for language lovers, although it also invites a person to immerse himself in the depths of knowledge in a pleasant and surprising way. The purpose of this type of work is to help readers. The thesaurus makes ideas more dynamic than a dictionary in the human mind by suggesting synonyms or related terms that are reflected in it, and further enhances opinions through easy ways of expression.

## REFERENCES

1. Shcherba, L. V. 1939. *Lexical Structure and Practical Application*.
2. Shcherba, L. V. 1974 *An Attempt at a General Theory of Lexicography*. <http://febveb.ru/feb/izvest/1940/03/403-089htm>.
3. *National Encyclopedia of Uzbekistan*. 2002. Tashkent: UzME State Publishing House.
4. *National Encyclopedia of Uzbekistan*. 2002. Tashkent: State Publishing House of the Republic of Uzbekistan.
5. Khojiev, A. 2002. *Explanatory Dictionary of Linguistic Terms*. Tashkent: State Publishing House of the Uzbek Academy of Sciences.
6. Osokina Svetlana Anatolevna. 2015. *Foundations of the Linguistic Theory of Thesaurus*. Barnaul.
7. Morkovkin, V. V. 1970. On the issue of educational thematic dictionary. *Russian Language and National School*, 2, 76-80.
8. Casaresa, X. 1958. *Introduction to Modern Lexicography*.
9. Vinogradova, B. B. 1977. *Lexicology and Lexicography*.
10. Lukashevich, N. V. 2011. *Thesauri in Information Retrieval Tasks*.
11. Abdurakhmonova, N. 2021. Formal-functional models of the uzbek electron corpus. *ANGLISTICUM. Journal of the Association-Institute for English Language and American Studies*, 10/8, 59-66.

**BABAZARIPOVA MALAKHAT ERGASH KIZI**

PHD STUDENT, NAVOI STATE UNIVERSITY,

UZBEKISTAN.

E-MAIL: <BOBOZARIPOVAMALOHAT555@GMAIL.COM>