

THE PRINCIPLE OF SELECTING WORDS BY CATEGORY IN THE ELECTRONIC FORM OF MATHEMATICAL DICTIONARIES

Batrova Adiba Erkinovna*

*Associate Professor,
Navoi State Pedagogical Institute,
UZBEKISTAN

Email id: botirova@gmail.com

DOI: **10.5958/2249-7137.2022.00381.0**

ABSTRACT

The article discusses some international experience in creating educational dictionaries, which contains the composition, main content and form of the dictionary, which should be taken into account when compiling it. The advantages of grammatical dictionaries in electronic form are revealed. Linguistic research has raised the issue of creating dictionary types by scanning large volumes of dictionaries to save time when working in a particular field, and classifying dictionaries from the point of view of the grammatical system within language levels to avoid wasting time.

KEYWORDS: *Dictionary, Mathematical Dictionaries, Electronic Dictionaries, Corpus, National Language, Corpus Creation, The Role Of Mathematical Dictionaries In The National Corpus, Foreign Experience, Linguistic Dictionaries*

INTRODUCTION

Since the emergence of the lexicographic educational lexicon network in the linguistic system, experts have been interested in what kind of dictionary is a dictionary, what are the criteria for its creation, what are its main functions, how it differs from the general dictionary, what are its advantages over other types of dictionaries. In world linguistics, as early as the beginning of the twentieth century, these questions were answered in accordance with the requirements of the time. In many developed countries, educational lexicography has left behind several stages of its development, and today is engaged in the theory and practice of creating a new genre of anthropocentric educational dictionary of the modern type. [1]

Decree of the President of the Republic of Uzbekistan dated October 21, 2019 No PF-5850 "On measures to radically increase the prestige and status of the Uzbek language as the state language" and other relevant regulations serves to a certain extent. In addition, the Decree PF-6084 of October 20, 2020 "On measures to further develop the Uzbek language and improve language policy in the country" sets the task to create linguistic and sectoral dictionaries, which will contribute to its implementation.

In this sense, the head of state said, "Each of us should consider the attention to the state language as attention to independence, respect and loyalty to the state language as respect and loyalty to the motherland, and make such a view a rule of our lives." [2], For us, the idea of

improving the conduct of the state language, the creation of sectoral and grammatical dictionaries also serves as a guide.

Literature review. Professor B. Mengliev identifies the following main functions of the dictionary, which have important social, political and cultural significance for the development of society:

- mother tongue and other language teaching;
- description and standardization of the native language;
- ensuring language and intercultural communication;
- Scientific research and interpretation of language vocabulary [3].

V.V. Morkovkin also draws attention to the fact that the purpose of the dictionary is to have a lingvodidactic character, that is, to be educational. Based on this, he emphasizes that educational lexicography is a science that emerged at the intersection of a separate linguometric direction, traditional lexicography and linguomethodology. In his opinion, the main sections of educational lexicography are:

- a) theory and practice of compiling a dictionary;
- b) the theory and practice of creating a lexical minimum;
- c) theory and practice of educational lexicostatistics;
- g) theory and practice of creating dictionary-type textbooks on lexicon;
- d) the theory and practice of semantics of language vocabulary should be presented in the dictionary attached to the textbook and given for a specific topic [4]

Despite the rigidity of the content device, the internal, microstructure of the dictionary can be variable not only in the textbook but also in the general dictionary. In particular, grammatical assignments, exercises, role-playing tests, and an audio message in an electronic dictionary can be added to the comment section of the textbook. In the modern textbook, additional information is added in tabular form [5].

B.Bahriddinova concludes that it is expedient to have a mega-structure of a modern educational dictionary in the following order:

- 1) access;
- 2) educational and methodical support;
- 3) dictionary body;
- 4) from the application;

macrostructure:

- 1) alphabetically;
- 2) thematic;
- 3) thematic-alphabetical dictionary from an article or dictionary;

microstructure:

- 1) keyword;
- 2) comment;
- 3) lexicographic pometa;
- 4) consists of illustrations.

The use of this criterion in the creation of grammatical dictionaries also allows you to easily place grammatical dictionaries in the educational building.

Modern lexicography has also activated a number of concepts such as computer lexicography, the concept of computer lexicography, electronic dictionary, the difference between electronic and traditional dictionaries, the structure of dictionary writing, types of electronic dictionaries, the advantages of electronic dictionaries and the prospects of computer lexicography. The concept of computer lexicography. It is a branch of computer lexicography. Lexicographic work includes the development of computer dictionaries, linguistic databases and support programs aimed at creating applied linguistics. The main tasks of traditional and computer lexicography are to determine the structure of the dictionary and the zones of access to the dictionary, as well as to develop principles for compiling dictionaries of different types. A dictionary is traditionally described as a collection of words organized with comments describing the structure and / or performance characteristics of these words.

Electronic dictionary. The difference between electronic and traditional dictionaries. An electronic (automatic, computer) dictionary is a set of words in a special computer format (e.g., machine translation systems) that are designed for human use or are an integral part of more complex computer programs. Accordingly, automatic dictionaries for automatic end-user dictionaries and text processing programs are distinguished.

Automatic dictionaries for the end user are often computer versions of well-known traditional dictionaries, such as: Oxford English Dictionary [6] (World War II). records are sorted by fields (see selecting all adjectives), performing an automatic search for all vocabularies that have a specific semantic component in interpretation, and so on.

Automatic dictionaries for machine translation systems, automatic abstraction, data acquisition and more. Dictionary writing differs significantly from others in terms of interface and structure. Their structural features, the scope of the dictionary are determined by the programs that interact with them. Such a dictionary can contain hundreds of dictionary access zones at once. The areas of lexicographic description are also very diverse: morphological, lexical, syntactic, semantic, and so on. The structure of a traditional dictionary usually includes the following components:

- Explain the principles of using the introductory dictionary and provide information about the structure of the dictionary entry;
- lexical units: a dictionary containing morphemes, lexemes, word forms or word combinations; each such unit having an appropriate annotation is a dictionary entry;
- indicators (indices);
- list of sources;

- lighting and alphabetical list.

In electronic dictionaries of these components, perhaps only the dictionary is mandatory, online dictionaries often have an alphabet with bridges behind each letter leading to the text of the dictionary entry. Almost every electronic dictionary available on disk (offline dictionary) or on the Internet (online dictionary) has an automatic search function, which allows you to significantly save the user's effort when working with the dictionary. [7]

The difference between grammatical electronic dictionaries and paper dictionaries also applies to their multimedia and hypertextuality: these features are expressed to a much greater extent than those printed in electronic dictionaries. Thus, hyperlinks can be placed behind any element of the dictionary entry or an item in the dictionary application menu. This gives the user additional options for searching and quickly navigating the required dictionary information, allowing you to find synonyms and antonyms for a particular word, words in the same semantic group, declension and conjugation paradigms, and more.

It makes it easy to link different dictionaries together so that online or offline dictionaries eventually turn into dictionary collections or portals. For example, after receiving the necessary information about the meaning of a word, the user can go to the comments of this word in other dictionaries and click on the link to know the specifics of its interpretation in special fields of knowledge (terminological dictionaries) or to get additional linguistic information about its form.

As a result, the structure of the electronic dictionary differs significantly from the structure of the printed dictionary, although the dictionary, which is the main part of the dictionary-dictionary articles, continues to form the core of the dictionary in both cases. [8]

3. Dictionary structure. The structure of a dictionary entry is typical and usually includes the following areas of dictionary writing that are appropriate for both traditional and computer lexicography:

- lexical input (vocabulary, lemma);
- grammatical information field;
- methodical garbage area; -
- value zone;
- zone of phraseological units;
- zone of etymology;
- sample zone and sample source.

However, it is possible to distinguish dictionary input zones, mandatory and optional zones for all dictionary units. For different types of dictionaries, the mandatory zone of dictionary entry is only lexical input, all other zones depend on the type of dictionary: for example, a semantic zone is necessary for an annotated dictionary, and for an orthoepic it is optional. The phraseological zone does not exist in the explanations of words not used in stable compounds, and the existence of an example zone and its source depends on the principles underlying the creation of the dictionary.

The number of dictionary input zones of a computer dictionary usually exceeds the number of dictionary input zones of a “paper” dictionary, which is due to significant memory resources and high speed of digital information processing by modern computers. But the amount of vocabulary information offered should be appropriate to the type of vocabulary: if the reader needs pronunciation, then “additional” information about the translation of the word being tested or its contextual meanings will only interfere with the user.

4. Types of electronic dictionaries. The classification of computer dictionaries can be done on the same principles as the classification of traditional dictionaries. Traditionally, linguistic, encyclopedic, and intermediate (linguistic and terminological) dictionaries are distinguished. Linguistic dictionaries describe the words themselves - their meanings, usage features, structural features, compatibility, connection with lexical systems of other languages, and so on. Encyclopedic dictionaries describe concepts, facts, and realities of the world around them, i.e. extralinguistic information. Intermediate types of dictionaries include linguistic and extralinguistic information. [9]

There are several types of linguistic dictionaries:

- interpretation of the meanings of words and their use in speech, including descriptive and normative dictionaries, which can be general and specific, for example, phraseological dictionaries, dictionaries of foreign words, etc.;

- Thesa-Uzbek dictionaries, which differ in the arrangement of the vocabulary subordinated to the thematic principle, not the alphabet, for example, include the semantic field "departure, departure, flight" of the Uzbek idiom in the category "movement", the semantic field to the category "time" placed, etc.;

B. B. Bahridinova "In the world educational dictionary, based on the age of the user, for preschool children" Learn to pronounce correctly ", " Speak correctly ", for primary school" Write correctly ", " Learn to pronounce correctly ", for high school" Spelling dictionary ", " Pronunciation textbooks ”have been created.

Mastering the norm is not an easy process. The reason is that the norm is a changing category, and the norm of language, like language itself, changes and evolves in relation to social development. This makes it difficult to create a normative-type textbook. Even the most developed countries in the world have problems with this. ”⁷ emphasizes

- Associative dictionaries, which are areas of associative relations in the dictionary; The dictionary of such a dictionary includes a list of frequencies and alphabetically ordered reactions obtained in lexeme-stimulus and psycholinguistic experiment, for example: “Dictionary of associative words of modern Uzbek language”;

- historical and etymological dictionaries that provide information about the history of words for a certain period from a certain date, indicating the emergence of new words and meanings, their disappearance and change, or explaining the origin of words;

- Dictionaries of language forms that correct word properties and do not interpret or play a supporting role, such as orthographic and orthoepic, word formation and morphemic (show how words are made up of morphemes and inventory them), grammatical (information about each word is which allows you to create a grammatically correct form), inverted dictionaries (words

are arranged in alphabetical order, with initial and not final letters, and on the right, not on the left: coat of arms, serb, damage, dumb, oak);

- Dictionaries of speech: dictionaries of difficulties and phrases;
- onomastics: anthroponymic dictionaries and toponymic dictionaries;
- Non-traditional, atypical linguistic objects that are subject to the description of the dictionary, for example, a dictionary of Uzbek political metaphors and dictionaries of poetic metaphors, epithets, author's dictionaries and dictionaries of conciliation.
- National Encyclopedia of Uzbekistan, Annotated Dictionary of the Uzbek Language.
- Official electronic dictionaries Uzbek spelling dictionaries (<http://savodxon.uz>) and English) languages.

There is also a large collection of various dictionaries on CDs and on the Internet provided by Duden Publishing (German, World War II. Duden. De) and Larousse (French, World War II. Larousse.az) and other dictionaries.

Advantages of electronic dictionaries. Computer dictionaries are usually created on the basis of text corporations using automatic processing and search of dictionary units. To do this, special programs - databases, computer card files, text processing programs allow you to automatically create dictionary entries, save and process dictionary data. Thus, creating an electronic dictionary involves the following steps:

- 1) form a body of texts and create a dictionary in parallel;
- 2) automatic formation of the sample body;
- 3) writing dictionary entries;
- 4) entering dictionary entries into the database (DB);
- 5) editing dictionary entries in the database;
- 6) text editing in the database;
- 7) create a dictionary text and form the original layout;
- 8) publishing a dictionary.

Of course, the above description of the process of creating an electronic dictionary can be adjusted depending on its type, research principles, and other factors. Comments of the creators of the electronic historical dictionary of the Uzbek language. But in any case, the use of computers and ready-made text corpus in computer lexicography allows to reduce the number of steps in the process of creating an electronic dictionary and save time in almost each of them.

This means that databases are used in computer lexicography instead of creating a dictionary map. Database entries allow you to automatically sort an array by selected parameters, select the desired instances, group them into groups, and so on. There is almost no specialized software shell on the market for lexicographic purposes. Modern databases such as ACCESS or PARADOX are very suitable for this purpose. To search for examples, dictionary creators can use computer software to build agreements, such as DIALEX. Publishing systems such as page

builder or word, which allow you to link styles to the dictionary access zones, alphabetization, indexing, etc., to create the original order (order) of dictionaries.

Probably the only example of a specialized computer program designed for computer lexicographic works is the “automated dictionary creation and processing program” (authors :). This program is very actively used in philological research and is described in detail in the textbook.

- Electronic dictionaries have positive aspects not only in the process of creation, but also in the process of use. In particular, the following advantages of using electronic dictionaries are noted:
 - 1) use of multimedia tools. Electronic dictionaries allow you to present the content of a dictionary entry in a variety of ways (different “projections” of the dictionary), including using a variety of graphical and multimedia tools not used in traditional dictionaries;
 - 2) Use of modern technologies. The data presented reflect various technologies of computational linguistics such as morphological and syntactic analysis, sound detection and synthesis;
 - 3) Convenient search. It will be possible to quickly retrieve information that is located anywhere in the dictionary and directly responds to a query created by the user in a form that is convenient for him. There is no need to remember the word clearly, the program itself offers options for the first letters.
 - 4) Relevance and dynamism. The electronic dictionary allows you to respond quickly to changes in language and the world, and it does not take much time and effort to release each subsequent version or make changes to the online version.
 - 5) Multi-word. Many electronic dictionaries have a terminological base that goes beyond the base of paper dictionaries and provides easier access to data using bridges.
 - 6) Variability in usage - the ability to use dictionaries on local and global networks, i.e. offline and online versions.
 - 7) Versatility - as a rule, programs allow you to work with multiple languages and translation directions at the same time. Any of the languages included in the dictionary can be used as an introduction.
 - Although there are many advantages to using electronic dictionaries, some of the problems that are relevant to traditional and computer lexicography remain unresolved.
 - Dictionaries should reflect the concept of lexical function, which allows to systematically describe the non-free combination of words described by the following examples of the Uzbek language: "war is underway" and "examination", "theories are advanced" and "existing ideas are presented", etc. .
 - The problem of describing semantics and the practical execution of grammatical deviations and word formation is not reflected in mass lexicographic practice. Each language has its own grammatical coding methods that are not systematically described in popular dictionaries.
-

- Dictionaries also do not have a system of concepts that can convey syntactic information to the average reader. The solution to this problem could be integrated dictionary descriptions based on formal models that take into account progressive lexicographic ideas. Dictionary access technologies should be organized on the same models.
- These problems can be solved in collaboration with theoretical and practical lexicographers, and computer tools will undoubtedly facilitate routine work on performing monotonous lexicographic operations.
- In general, we emphasize that computer lexicography aimed at creating electronic dictionaries is a very promising and necessary direction of computational linguistics, because the products created by it - electronic dictionaries - are characterized by versatility, multimedia, integration of latest technological solutions, material relevance and user satisfaction. arrange access to the required information.
- Analysis and results. Issues of developing the structure of electronic grammar dictionaries (for Uzbek, German, Spanish, English) and their functions, as well as some aspects are the construction of a user interface. Dictionaries are information processing systems (in morphological analysis and text synthesis algorithms) designed for use in information correction systems, as well as for use in automatic text contours. In developing the content of the dictionary, the requirements for this type of software products are taken into account.
- Lexicographic systems are defined
- successful use of information;
- The theory of lexicographic systems is an integral lexicographic system;
- combining the functions of language flexion, orthoepy, phraseology, synonymy and antonymy;
- use as a finished lexicographic product.

Electronic grammatical dictionaries are mainly intended for use as a means of automatic morphological analysis in the MP system (text marking, lemmatization and synthesis in morphological stages). In addition, the dictionary should be available to the user, a reference tool (word search, word units that provide information about the exact inflection). The structure of the dictionary also has certain requirements for such features of the purpose (the linguistic information provided in the dictionary should be sufficient for the following. Perform all necessary functions; different methods of access to this information should be provided). At different stages of the texts it will be necessary to obtain the original form of the word with MP. Text (when analyzing input text) and synthesizing the desired word form from the original.

Dictionary A. The source of linguistic information about A. Zaliznyak (hereinafter in the text) is in Uzbek. The technology of creating a computer database of the Uzbek grammar dictionary includes:

- electronic translation (scanning) of paper;
- edit scanned text;

- Development of the structure of the lexicographic system, the language and organization of its definition
- identifiers of structural elements;
- electronic text to the lexicographic database (LBD)
- automatic rotation;
- in accordance with the developed structure;
- Development of paradigmatic classification algorithms and their software;
- formation of paradigmatic (automatic indexing of lexemes);
- grammar class codes and paradigmatic class numbers).

Grammars and dictionaries of German, English, and Spanish have been used as a source of linguistic information to compile grammatical dictionaries.

Principles of presentation of lexical material (Data structure) At the internal level of the architecture of the lexicographic system, the linguistic structure of the Uzbek language is represented by a model of data relations, the relationship of which is represented by tables:

- Table of names of registry units and the number of paradigmatic classes (type field), indicating the code of the part of the lexico-grammatical class;
- flexible quasi-rotation table for each grammatical form (field);
- quasi-flex reflexes of each paradigmatic class (type field) are given;
- paradigmatic intendaent table of classes showing parameters and properties that are the same for each;
- Table of parts of lexical and grammatical classes and their codes;
- table of inflection types.

The relationship between the tables is done by the number of paradigmatic classes (type field); and between the tables, the title, the parts are given.

In developing the data structure for other languages (English, Spanish, German), an approach similar to that described above was used to consolidate their representations, i.e., the unit of the main dictionary was quasi-newly selected. The only difference is that in the LBD of these languages, in addition to the paradigm deviations in the description, the type of procedure (analytical) used in the construction of specific forms is also included.

The quasi-bases of regular verbs in English and Spanish, most nouns and adjectives correspond to the original form of the word. The quasi-lexical representation of a lexical unit is mainly based on only irregular verbs and a small number of nouns.

German is classified as a synthetic-analytical type of language. Simple (synthetic) forms are formed in it by the flexion method, while complex (analytical) forms are formed according to certain schemes (procedural). It is also characterized by a large number of substitutions in the word base and the phenomenon of separation of prefixes in a particular group of verbs. All of

these features require that the processes of the German language take them into account, first and foremost, in the separation of the flexion set.

Hence, language units need to make some additions to the paradigmatic classes, secondly, to the data structure. In addition to the tables shown in the figure, the data structure includes a table. A table showing the types of prefixes to be allocated, as well as the types of procedures for constructing analytic forms. Additional fields are included in the indent table, indicating the type of substitutions, as well as the delimiter of the prefix⁸

CONCLUSION/RECOMMEND

In conclusion, there is a need to develop separate grammatical dictionaries for noun and verb phrases, as the noun phrase itself is studied in the largest number of hours in the general secondary education system. The student will not be able to carry a general commentary or spelling dictionary for each lesson. As any innovation goes towards convenience, it acquires originality, usefulness, especially when it takes a long time to find a word from a spelling dictionary. Paper dictionaries take up too much time, and even electronic dictionaries do not fit the research topic, so for statistical analysis it is necessary to separate words from another word group and draw a conclusion. In this sense, in the construction of grammatical dictionaries, the criterion for classification should be the division into word groups. At the same time as corpus linguistics is evolving, it is necessary to create a separate dictionary (in electronic form) of word groups for a rapid analysis of both the frequency of use of linguistic units and the definition of primary or secondary meanings. In addition to being extremely necessary for this linguistic research, dictionaries of nouns and verbs, as well as other categories, can also be a handy guide in the educational process. In particular, the Uzbek language is taught in depth in the President's creative schools, which helps to acquaint students with the grammatical structure of the language.

REFERENCES:

1. Resolution of the President of the Republic of Uzbekistan Sh.M.Mirziyoev No. PP-4479 "On the broad celebration of the thirtieth anniversary of the adoption of the Law of the Republic of Uzbekistan" On the State Language ". Available at: uza.uz.
2. Bahridinova B. Educational dictionary in Uzbekistan: linguistic bases, history and prospects. Doctors of Philological Sciences. dissertation. Samarkand. 2020. 23p.
3. Sayfullaeva RR, Mengliev BR, Boqieva GH, Qurbonova MM, Yunusova ZQ, Abuzalova MQ. Modern Uzbek literary language. Study guide. Tashkent, 2006. 391p. B.125
4. Morkovkin VV. Educational lexicography as a special linguo-methodological discipline. Actual problems of educational lexicography. Moscow: Russian language, 1977, pp. 28-37.
5. Gerd AS, Sergeyev VN. To the definition of the concept of a dictionary. Problems of lexicography: Collection of Art. St. Petersburg: Publishing House of St. Petersburg. un-ta, 1997. pp. 191-203.
6. Hojiev A. Interpretation of linguistic terms. Tashkent: UzMEDavlatilmiynashriyoti, 2002. 105p.

7. Explanatory dictionary of the Uzbek language. Tashkent: "Uzbekistan Publishing House" State Unitary Enterprise, 2020. 680 p.
8. Lyubchenko TP. Ukrainian Language and Information Foundation of the National Academy of Sciences of Ukraine.
9. Ozhegov Dictionary online. Available at: <http://slovarozhegova.ru>.