

Reformatskiy A. Myсли o terminologii [Thoughts about Terminology] / A. Reformatskiy // *Sovremennyye problemy terminologii v nauke i tehnike: Sb. st.* — М. : Nauka, 1969. — P. 163-199 [in Russian].

8. Романова О.О. Спеціальна лексика української мови як об'єкт лінгвістичного дослідження: термін і номен // *Термінологічний вісник.* — 2011 — Вип.1 — С.55-62.

Romanova O.O. Spetsialna leksyka ukrainskoi movy yak ob'iekt lnhvistychnoho doslidzhennia: termin i nomen [Special Vocabulary of the Ukrainian Language as an Object of Linguistic Research: Term and Nomenclature Sign] // *Terminolohichniy visnyk.* — 2011 — vol. 1 — P.55-62 [in Ukrainian].

9. Українська мова: Енциклопедія / Редкол. Русанівський В.М., Тараненко О.О., Зяблюк М.П. та ін. — 2-ге вид., випр. і доп. — К. : "Українська енциклопедія" ім. М.П. Бажана, 2004. — 824 с.

Ukrainska mova: Entsyklopediia [Ukrainian Language: Encyclopedia] / eds. Rusanivskiy V.M., Taranenko O.O., Ziabliuk M.P. and others. — 2nd edition, corrected and supplemented. — К. : "Ukrainska entsyklopediia" named after M.P. Bazhan, 2004. — 824 p.[in Ukrainian].

10. Щерба Д. В. Термін та його дефініція як головні онтологічні поняття термінознавства / Д. В. Щерба // *Вісн. Житомир. держ. ун-ту ім. І. Франка.* — 2006. — Вип. 28. — С. 237-239.

Shcherba D. V. Termin ta yoho defynitsiia yak holovni ontolohichni poniattia terminoznavstva [Term and its Definition as the Main Ontological Concepts of Terminology] / D. V. Shcherba // *Visn. Zhytomyr. derzh. un-tu im. I. Franka.* — 2006. — vol. 28. — P. 237-239 [in Ukrainian].

*Alina Teletska*

*Vasyl' Stus Donetsk National University*

*Vynnytsia*

*Research Supervisor: V. I. Kalinichenko, PhD in Philology, Ass. Prof*

*Language Supervisor: N.V.Striuk, Lecturer*

## **TYPES OF LANGUAGE CORPORA AND THEIR USE**

**Introduction.** Nowadays, there are many studies, which have to be confirmed with certain official data. Language corpora has acquired wide popularity and use in this area. It is able not only to solve linguistic problems but it can be applied to other sciences. **The objective of the paper** is to define the language corpora, to reveal its types and applications.

*Corpus linguistics* has several meanings. The first one defines it as a method of carrying out linguistic analyses. As it can be used for the investigation of many kinds of linguistic questions and as it has been shown to have the potential to yield highly interesting, fundamental, and often surprising new insights about language, it has become one of the most wide-spread methods of linguistic investigation in recent

years. The second meaning: corpus linguistics is the analysis of naturally occurring language on the basis of computerized corpora. Usually, the analysis is performed with the help of the computer, i.e. with specialised software, and takes into account the frequency of the phenomena investigated.

A corpus is a systematic collection of naturally occurring texts (of both written and spoken language). Which means that the structure and contents of the corpus follow certain extralinguistic principles and information on the exact composition of the corpus is available to the researcher (including the number of words in each category and in the whole corpus, how the texts included in the corpus were sampled etc.) [1: 2].

The aim of corpus linguistics is to improve language description and theory. Corpus data are essential for accurately describing language use, and have shown how lexis, grammar, and semantics interact. This in turn has applications in language teaching, translation, forensic linguistics, and broader cultural analysis [2: 106].

*Types of text corpora.* A text corpus can be classified into various categories by the source of the content, metadata, the presence of multimedia or its relation to other corpora. The same corpus can fall into more than one category if it fulfils the criteria for more categories:

- *a monolingual corpus* is the most frequent type of corpus. It contains texts in one language only. The corpus is usually tagged for parts of speech and is used by the wide range of users for various tasks from highly practical ones, e.g. checking the correct usage of a word or looking up the most natural word combinations, to scientific use, e.g. identifying frequent patterns or new trends in language.

- *a parallel corpus* consists of two monolingual corpora. One corpus is the translation of the other. The user can search for all examples of a word or phrase in one language and the results will be displayed together with the corresponding sentences in the other language. The user can then observe how the search word or phrase is translated.

- *a multilingual corpus* is very similar to a parallel corpus. The two terms are often used interchangeably. A multilingual corpus contains texts in several languages which are all translations of the same text. A user can also decide to work with one language to use it as a monolingual corpus.

- *a comparable corpus* is a set of two or more monolingual corpora whose texts relate to the same topic., however, they are not translations of each other, and therefore, they are not aligned. When users search these corpora, they can use the fact, that the corpora also have the same metadata.

- *a learner corpus* is a corpus of texts produced by learners of a language. The corpus is used to study the mistakes and problems learners have when learning a foreign language.

- *a diachronic corpus* is a corpus containing texts from different periods and is used to study the development or change in language.

- *a specialized corpus* contains texts limited to one or more subject areas, domains, topics etc. Such corpus is used to study how the specialized language is used.

- *a multimedia corpus* contains texts which are enhanced with audio or visual materials or other type of multimedia content [3].

*A brief history.* The first computer-readable corpora, compiled in the 1960s, are very small by contemporary standards, but still useful because of their careful design. The Brown corpus (from Brown University in the USA) is one million words of written American English, sampled from texts published in 1961: both informative prose, from different text-types (e.g., press and academic writing), and different topics (e.g., religion and hobbies); and imaginative prose (e.g., detective fiction and romance). Parallel corpora were designed to enable comparative research: the LOB corpus (from the universities of Lancaster, Oslo, & Bergen) contains British data from 1961; Frown and FLOB (from Freiburg University, Germany) contain American and British data from 1991; and ICE (International Corpora of English) contains regional varieties of English, such as Indian and Australian. Similar design principles underlie the Lund corpus of spoken British English (from University College London and Lund University), which contains around half a million words, divided into samples of the usage of adult, educated, professional people, including face-to-face and telephone conversations, lectures and discussions. By the late 1990s, some corpora consisted of hundreds of millions of words. The Bank of English (at COBUILD in Birmingham, UK) and the British National Corpus (BNC) had commercial backing from publishers, who have used the corpora to produce dictionaries and grammars. The 100-million-word BNC is also carefully designed to include demographically and stylistically defined samples of written and spoken language. The Bank of English arguably over-emphasizes mass media texts, but these are very influential, and it still has a range of text-types and advantages of size: over 400 million words by 2001 [2: 111-112].

Corpora are just sources of evidence, available to all linguists, theoretical or applied. A sociolinguist might use a corpus of audio-recorded conversations to study relations between social class and accent; a psycholinguist might use the same corpus to study slips of the tongue; and a lexicographer might be interested in the frequency of different phrases. The study might be purely descriptive: a grammarian might want to know which constructions are frequent in casual spoken language but rare in formal written language. Alternatively, it might have practical aims: someone writing teaching materials might use a specialized corpus to discover which grammatical constructions occur in academic research articles; and a forensic linguist might want to study norms of language use, in order to estimate the likelihood that linguistic patterns in an anonymous letter are evidence of authorship [2: 107].

**Conclusion.** There are different language corpuses, which are appropriate for special aims and audience. A text corpus helps to conduct fast searching in a big quantity of data on a wide range of subjects. It is a valuable resource of learning information and it provides high quality of research works. Nowadays, language

corpora has reached great achievements, but there are also tasks, which have to be studied and solved in order to open more opportunities.

### References

1. Nesselhauf N. Learner Corpora and their Potential for Language Teaching / N. Nesselhauf, J. Sinclair // *How to Use Corpora in Language Teaching* / N. Nesselhauf, J. Sinclair. – Amsterdam & Philadelphia: Benjamins, 2004. – P. 125–152.
2. Davies A. *The Handbook of Applied Linguistics*. / A. Davies, C. Elder. – Wiley: Language Arts & Disciplines, 2006. – 866 p.
3. Types of Text Corpora [Electronic resource]. – Retrieved from: <https://www.sketchengine.co.uk/user-guide/user-manual/corpora/corpus-types/>

*Yevhenia Yumadilova*

*Vasyl' Stus Donetsk National University*

*Vinnitsia*

*Research Supervisor: L.M. Koval, Doc. of Philology, Full.Prof.*

*Language Striuk: Yu. Ya. Khankishyieva, Lecturer*

## CHARACTERISTICITY OF MEN'S NAME IN THE STORAGE OF STEPANIVKA OF THE SHAKHTAR DISTRICT OF DONETSK REGION

**Introduction.** For a long time proper names have attracted and still continue to attract the attention of linguists. Nowadays, the most thoroughly researched area is surnames, which has been examined by Yu. K. Redka, O. D. Nediilka, I. D. Sukhomlyn, V. D. Poznanska, G. Ye. Buchko, etc. Also, quite deeply this area is learnt by O. V. Antoniuk, Yu. O. Nalyvaychenko, N. M. Brener, M. V. Tyminsky. Personal names of people have also repeatedly been the object of specific studies. Particularly it was researched on the material of historical monuments studied by R. Y. Kerst, M. I. Seniv, P. I. Ostash, M. O. Demchuk, M. L. Hudash, I. M. Zheleznyak, L. L. Humetska; modern personal names of Ukrainians are studied by L. O. Belei, I. D. Sukhomlyn, T. V. Bug, L. O. Kravchenko, P. P. Chuchka, L. T. Masenko, etc. Despite the numerous profound researches of Ukrainian onomasts, many anthroponymy problems remain open. Among them is the problem of anthroponymy creation of separate populated settlements.

That is, **the purpose of this paper** is to describe an anthroponymy portrait of the village Stepanivka of Shahtarskiy district of Donetsk region.

Common feature to the whole Ukraine is that name's system has its own peculiarities in certain territories. It depends on the situation of the speech; the names can acquire a variety of modifications. Except official and full documentary homonymies, the interviewers fix the different colloquial everyday life variants (abbreviation or truncated, chopped-caressing and rude-disparaging).