

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ**  
**«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»**  
**ФАКУЛЬТЕТ ЛІНГВІСТИКИ**  
**КАФЕДРА ТЕОРІЇ, ПРАКТИКИ ТА ПЕРЕКЛАДУ АНГЛІЙСЬКОЇ МОВИ**

«На правах рукопису»  
УДК \_\_\_\_\_

«До захисту допущено»

Завідувач кафедри

\_\_\_\_\_  
(підпис) (ініціали, прізвище)

“ \_\_\_ ” \_\_\_\_\_ 20\_\_ р.

## Магістерська дисертація

на здобуття ступеня магістра

зі спеціальності \_\_\_\_\_ 8.02030304 “Переклад” \_\_\_\_\_  
(код і назва)

на тему: \_\_\_\_\_ **Лінгвістичний корпус сфери інформаційних технологій:  
перекладознавчий та лексикографічний аспекти** \_\_\_\_\_

Виконав (-ла): студент (-ка)  5  курсу, групи  ЛА-51м   
(шифр групи)

\_\_\_\_\_  Бондарець Ольга Михайлівна  \_\_\_\_\_  
(прізвище, ім'я, по батькові) (підпис)

Науковий керівник  Демиденко О.П., кандидат пед. наук, доцент  \_\_\_\_\_  
(посада, науковий ступінь, вчене звання, прізвище та ініціали) (підпис)

Консультант \_\_\_\_\_ \_\_\_\_\_  
(назва розділу) (науковий ступінь, вчене звання, прізвище, ініціали) (підпис)

Рецензент  Туришева О.О., кандидат філологічних наук, доцент  \_\_\_\_\_  
(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали) (підпис)

Засвідчую, що у цій магістерській дисертації немає запозичень з праць інших авторів без відповідних посилань.

Студент \_\_\_\_\_  
(підпис)

Київ – 2016 року

## Реферат

Технічна документація сфери інформаційних технологій використовується для архітектурних рішень у розробці продукції. У зв'язку з швидким збільшенням аутсорсингових компаній, які спрямовані на англомовний ринок, виникає необхідність якісного забезпечення продукції супроводжуваними документами іноземною мовою. Для того щоб визначити основні положення при створенні технічної документації, необхідно визначити основні вимоги до її написання на прикладі оригінальних англомовних документів та створити їх лінгвістичний корпус. Створений корпус значно полегшить процес аналізу та перекладу документів, оскільки з його допомогою можна прослідкувати функціонування мовних одиниць в їх природньому оточенні та зрозуміти контекст їх вживання.

**Актуальність** даної дослідницької роботи полягає у необхідності застосування інноваційних прикладних технологій для об'єктивізації та прискорення досліджень значних обсягів стандартних текстів та їх перекладу. У цьому контексті на особливу увагу заслуговують методи корпусної лінгвістики, які активно наразі використовуються у лексикографії, теорії та практиці перекладу.

Дослідженнями лінгвістичних корпусів займалися такі вчені, як Баранов А. Н., Городецький Б. Ю., Казакевич О. А., Риков В. В. та інші.

**Метою** даної роботи є дослідження специфіки перекладу англомовних інструкцій як технічної документації сфери інформаційних технологій українською мовою із застосуванням прикладних лінгвістичних технологій.

У процесі дослідження ми ставимо перед собою наступні завдання:

- визначити основні положення корпусної лінгвістики;
- розглянути лінгвістичний та інформаційний аналіз тексту;
- дослідити функціонування конкордансу у мовному середовищі;
- визначити типологічну стратифікацію технічної документації;
- ознайомитися з оптимальними системами автоматизованого та машинного перекладу;
- виокремити вид документації, дослідження якого є найбільш актуальним на даному етапі розвитку інформаційних технологій;

- здійснити лінгвістичний аналіз інструкцій;
- створити конкорданс на матеріалі технічної документації;
- проаналізувати функціонування мовних одиниць, їх сполучуваність у конкордансі;
- визначити трансформації, які використовуються для адекватного та еквівалентного перекладу технічної документації.

**Об'єктом** даного дослідження виступає англомовний лінгвістичний корпус сфери інформаційних технологій.

**Предметом** дослідження прийомі відтворення в україномовному перекладі складових лінгвістичного корпусу, створеного на базі англомовних інструкцій як стандартної документації сфери інформаційних технологій.

Специфіка теми роботи, завдань і матеріалу зумовила комплексний підхід до вибору методів дослідження: комунікативний, трансформаційний, порівняльний, структурний.

Магістерська дисертація складається зі вступу, трьох розділів, висновків, списку використаної літератури, що складає 79 лексикографічних джерел та додатків.

**Апробація** основних результатів дослідження здійснювалася за рахунок участі у міжнародній науково-практичній конференції «Сучасні наукові дослідження представників філологічних наук та їхній вплив на розвиток мови і літератури» (8-9 квітня, 2016, м. Львів) з тезами «Переклад документації сфери інформаційних технологій» та участі у міжнародній науково-практичній конференції «Нове у філології сучасного світу » (10 – 11 червня, 2016, м. Львів) з тезами «Лінгвістичний корпус документації сфери інформаційних технологій: перекладознавчий аспект ».

В процесі написання магістерської дисертації було створено лінгвістичний корпус на базі стандартної технічної документації та здійснено його аналіз.

Обрана нами інструкція Android Lollipop Quick Start Guide складається з 13 427 символів, 11 050 слів та 650 речень. За допомогою корпусу було доведено, що інструкція написана саме у науковому стилі. У тексті використовується певний набір мовно-стилістичних засобів: спеціальні слова (терміни), складні синтаксичні конструкції (у яких має місце суворо впорядкований зв'язок, наприклад, за рахунок

вставних конструкцій); речення, ускладнені узагальнюючими родовими найменуваннями.

Для перекладу документів сфери інформаційних технологій використовують автоматизований та машинний переклад. При перекладі інструкцій використовується лексичні трансформації, а саме калькування та транслітерація. Серед граматичних трансформацій, дослівний переклад є найбільш популярним. Серед лексико – семантичних трансформацій найчастіше використовується еквівалентна заміна.

Отже, лінгвістичний корпус стандартної технічної документації сфери інформаційних технологій є інноваційним напрямком дослідження, який використовується при аналізі та перекладі текстів, проте містить ще багато не досліджених аспектів.

Ключові слова: лінгвістичний корпус, технічна документація, керівництво користувача, науковий стиль.

### **Abstract**

Technical documentation of IT field is being used for architectural solutions in product development. Due to the rapid growth of outsourcing companies aimed at English-speaking market, it is necessary to provide high-quality documentation when promoting products. In order to determine the key points of technical documentation, there is a need to determine the main requirements at the example of original documents in English by creating its linguistic corpus. It facilitates the process of analysis and translation of documents and depicts language units functioning in their natural environment.

The **relevance** of this research lies in the need of applying innovative technologies to accelerate investigations of big volume texts and its translation. In this context, special attention deserves methods of corpus linguistics that are actively used in lexicography, theory and translation.

Linguistic corpuses were investigated by such scholars as A. Baranov, B. Gorodetsky, O. Kazakevich, V. Rykov and others.

The **aim** of this work is to study peculiarities of translating manuals to Ukrainian with the help of applied programs.

In the study, we set the following objectives:

- Identify the fundamental principles of corpus linguistics;
- Conduct the linguistic and informational text analysis;
- Examine the functioning of concordance in the language environment;
- Identify typological stratification of technical documentation;
- Compare the automated and machine translations;
- Distinguish the type of document that are of frequent use in the IT field;
- Conduct the linguistic analysis based on the manuals;
- Create the concordance of technical documentation;
- Analyze the functioning of language units, their compatibility with the help of concordance;
- Define the applicable transformations when translating technical documentation.

The **object** of investigation is a linguistic corpus of IT documents.

The **subject** of the work is to apply the components of linguistic corpus based on English manuals to Ukrainian documentation.

Specificity of topics, tasks and material led to a comprehensive approach to the choice of research methods: communicative, transformation, comparative, structural.

Master's thesis consists of an introduction, three chapters, conclusions, list of references that contains 79 lexicographical sources and additional materials.

**Approbation** of the research was carried out by the participation in the international research and practice conference "Modern scientific research conducted by philologists and their impact on the development of language and literature" (April 8-9, 2016, c. Lviv) with the paper "Translation of IT documents" and in the international scientific conference "New Studies in the modern world" (10 - 11 June 2016, c. Lviv) with the thesis "Linguistic corpus of IT documents, its translation aspect".

In the process of writing a master's thesis we created the linguistic corpus based on the standard IT documentation. We have analyzed the “Android Lollipop Quick Start Guide” manual that consists of 13 427 symbols, 11 050 words and 650 sentences. With the help of corpus, we can prove the scientific style of the manual. The text uses a set of linguistic and stylistic means: special words (terms), complex syntactic constructions (which is a strictly ordered connections, for example, with parenthesis); sentences complicated by generic names.

We were using automated and machine translation when translating IT documents. Speaking about lexical transformations, the most frequently used ones were loan translation and transliteration. Among grammatical transformations, literal translation was the most popular. Among the lexical - semantic transformations equivalent replacement is used in most cases.

All in all, the linguistic corpus of It documentation is an innovative area of research used in the analysis and translation of texts, but still contains many unexplored aspects.

**Keywords:** linguistic corpus, technical documentation, user guides, scientific style.

## **ЗМІСТ**

|   |    |
|---|----|
| <b>ВСТУП</b> .....  | 3  |
| <b>РОЗДІЛ 1 Корпусна лінгвістика як сучасний напрямок лінгвістичних досліджень</b> .....                  | 6  |
| 1.1 Функціональні характеристики лінгвістичних корпусів у аспекті міждисциплінарних досліджень мови ..... | 6  |
| 1.2 Конкорданс як сучасна технологія аналізу мовних даних.....  | 15 |
| 1.3 Специфіка інформаційної обробки тексту із застосуванням конкордансу.....                              | 26 |
| 1.4 Комп'ютерна лексикографія та її взаємозв'язок з корпусною лінгвістикою.....                           | 32 |
| <b>ВИСНОВКИ ДО РОЗДІЛУ 1</b> .....  | 35 |

|  |    |
|--|----|
| <b>РОЗДІЛ 2 Сучасні технології мовної обробки, аналізу та перекладу технічної документації</b> .....   | 37 |
| 2.1 Типологічна стратифікація стандартної документації.....  | 37 |
| 2.2 Сучасні методи обробки й аналізу технічної документації засобами інформаційних технологій.....   | 40 |
| 2.4 Порівняльна характеристика функціональної специфіки сучасних перекладацьких технологій.....  | 44 |
| 2.4.1 Системи автоматизованого перекладу (Translation Memory).....   | 45 |
| 2.4.2 Програми машинного перекладу (CAT).....  | 48 |
| <b>ВИСНОВКИ ДО РОЗДІЛУ 2</b> .....   | 53 |
| <b>РОЗДІЛ 3 Інструкція як стандартний документ та об'єкт прикладних лінгвістичних досліджень</b> .....   | 55 |
| 3.1 Функціонально-стилістичні характеристики інструкції як стандартного документа сфери ІТ.....  | 55 |
| 3.2 Лексично-граматичні характеристики англомовних інструкцій та керівництв користувача.....   | 57 |
| 3.3 Специфіка лінгвістичного аналізу англомовних інструкцій з використанням технологій Sketchengine: лексикографічний аспект.....  | 65 |
| 3.4 Перекладознавчий аналіз прийомів відтворення функціонально-стилістичних та лексико-граматичних характеристик англомовних інструкцій сфери ІТ в україномовному перекладі..... | 74 |
| <b>ВИСНОВКИ ДО РОЗДІЛУ 3</b> .....   | 79 |
| <b>ЗАГАЛЬНІ ВИСНОВКИ</b> .....   | 82 |
| <b>SUMMARY</b> .....   | 85 |
| <b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</b> .....  | 86 |
| <b>ДОДАТОК А</b> .....   | 93 |

