

## **ПАРАЛЕЛЬНИЙ КОРПУС ТЕКСТІВ: ТЕОРЕТИКО-МЕТОДОЛОГІЧНИЙ ТА ЛЕКСИКОГРАФІЧНИЙ АНАЛІЗ, ПРИНЦИПИ ЗАСТОСУВАННЯ**

У статті розглядаються теоретико-методологічні аспекти обробки двох чи більше текстів на основі корпусного перекладу. Вказується на позитивний лексикографічний рівень паралельного корпусного перекладу (збагачення тексту стійкими словосполученнями, ідіомами, уточненнями, термінами, термінологічними виразами тощо). Зроблено висновок про збереження в окремих конструкціях-колокаціях стійких лексичних одиниць, які піддаються автоматизації та графічній корекції.

**Ключові слова:** колокація, паралельний корпусний підхід, лексикографічний аналіз, автоматизація тексту, авторський переклад.

**Актуальність дослідження** тісно пов'язана з проблемою лексикографічної диференціації паралельного корпусу текстів офіційно-ділового стилю. Під час застосування окремих корпусних ресурсів (Британського національного корпусу, Браунівського корпусу, Національного корпусу української та російської мови, корпусу російських текстів СинТагРус, ПАРТЕКС та інших) виникає проблема деталізації українських термінів та фразеологічних одиниць колокації тексту. Вдосконалення прикладної системи паралельного корпусу аналізу тексту зумовлює необхідність досліджувати великі обсяги природно-мовного матеріалу. Тому дослідження ґрунтується на об'єктивних корпусних даних, що опрацьовуються достовірними методами.

**Мета дослідження** – визначити теоретичні та методологічні принципи застосування паралельного корпусу текстів, пояснити лексикографічні зміни окремих одиниць колокації тексту.

### **Концептуальні завдання:**

1) Обґрунтувати взаємозв'язки між корпусним підходом та паралельним корпусом тексту;

2) З'ясувати принципи побудови паралельного корпусу текстів різних функціональних стилів;

3) Визначити похибки паралельного корпусного підходу (вирівнювання, розбіжності, збігання тексту; лексико-граматичні та фразеологічні розбіжності).

**Виклад основного матеріалу, результати дослідження.** У структурі автоматизованої обробки тексту робота з корпусами (масивами текстів) стала одним із важливих напрямів сучасних лінгвістичних досліджень. Корпус текстів найчастіше класифікують за ступенем організації і структурованості (електронний архів, електронна бібліотека, корпус текстів, субкорпус); за хронологічною ознакою (синхронний, моніторний, діахронний); за індексацією (простий, анотований); за мовою (одномовний, двомовний, багатомовний); за способом застосування і використання корпусу (дослідницький, ілюстративний, паралельний); за способом існування корпусу (динамічний, статичний).

У процесі пошуку корпус може висвітлити можливі варіанти обраного слова в безпосередньому контексті, на підставі чого перекладач може визначитися з вибором синоніма при перекладі або із сполучуваністю слів (Риков В. В.) [8]; частотність уживання слова в певній галузі знань (Ягунова Е. В.) [9]; слова, які найчастіше стоять поруч з обраним словом (Касьяненко М. К.) [7, с. 16]; інтертекстуальність: значення слова зі суміжністю його вживання; розвиток концептів у часі (Лебедев К. М.) [7, с. 19].

Багатомовні корпуси можуть бути двох типів: вирівняні або паралельні (*aligned*) і не вирівняні (*not aligned*). «Вирівняність» означає, що в корпусі існує чіткий зв'язок між одиницями перекладу, які відповідають один одному: можна швидко знайти, як те чи інше слово перекладається на іншу мову. Зазвичай такими одиницями перекладу виступають так звані «пропозиції». До кожної фрази існує мовний відповідник та його еквіваленти іншою мовою.

Відповідно, сучасні дослідження паралельного корпусу текстів тісно переплітаються: 1) із категоріями корпусної лінгвістики, як можливості цифрової обробки текстів, їх лексико-граматичного аналізу, розробки систем машинного перекладу, з метою створення комп'ютерних словників (концепція Волченкової К. Н.); 2) із репрезентацією та розвитком корпусного підходу (концепція Д. Сінклера) [13, с. 86].

Відповідно до певного процесу автоматизованого перекладу застосовується паралельний корпус текстів та порівняльний корпусний підхід. Паралельний корпус текстів (*parallel/translational corpus*) трактується здебільшого, як перекладацький корпус на мові оригіналу та відповідних текстах мови перекладу. Яскравою характеристикою паралельного корпусу є розмітка, яка при великих за обсягом корпусах, виконується автоматичним способом.

Вихідна мова корпусу має відповідати пропозиції мови перекладу. Якщо порушують співвідношення, здійснюється ручне редагування, щоб не допустити розбіжності пропозицій мовою оригіналу з пропозиціями мовою перекладу. Такий корпус дає кілька можливостей: встановити відповідності між потрібними елементами корпусу; виявити їх послідовність; створити список подібних одиниць, що є синонімічно близькими. Різновидом паралельного корпусу є зворотний перекладацький корпус (*back-translational corpus*), у якому можна змінити напрямок перекладу «мова А» переходить в «мову Б», а «мова Б» у «мову А». Можна погодитися з М. Бакером [10, с. 230], який вважає, що на відміну від порівняльного корпусу, у якому складно виявити параметри зіставлення, паралельний корпус володіє точнішими відповідниками при перекладі.

Паралельний корпусний підхід цілком доцільно застосовувати в корпусних ресурсах: Британський національний корпус, Браунівський корпус, Національний корпус української та російської мови, корпус російських текстів СинТагРус тощо.

Клименко С. В., розглядаючи похідні моделі паралельного корпусу тексту, враховує (в процесі їх створення) роль логічних дедуктивно-індуктивних принципів [8]. Наприклад, більшість існуючих і новостворюваних паралельних корпусів за типом об'єкту, можна поділити на корпуси поширених мов (стандартні, спеціалізовані, паралельні) і корпуси малопоширених і вимираючих мов. Оскільки час складання корпусу для вимираючих мов є критичним фактором, пріоритети зміщуються в бік кількісних аспектів, і в ролі упорядника може виступати будь-яка людина. Таким чином їх наповнення зводиться до опису програмно-апаратних способів реєстрації мовних даних, окреслення принципів маркування текстів за частинами мови, принципів транскрипції і транслітерації усного мовлення.

До проблемних факторів паралельного корпусу текстів відносимо вирівнювання та сегментацію (Вероніс Й., Добровольський Д. О., Кретов А. А. і Шаров С. А.) [6; с. 282], колокацію відповідно до функціонального стилю (Бобкова Т. В., Перебийніс В., Сорокін В.) [1, с. 159], лексико-граматичні похибки кодування (Большакова Є. І., Борисова Є. Г., Митрофанова О. А.) [5, с. 82].

Вирівнювання текстів за пропозиціями створює складні завдання, часто з множинними рішеннями, що виникають у результаті неоднозначності рішення самого завдання сегментації тексту на пропозиції (багатозначності точки як знака пунктуації, особливостей опису прямої мови в текстах художньої літератури, публіцистики, офіційно-ділових документів, відсутності фіксації кінця пропозиції); у результаті розбіжності розподілу вхідного і вихідного текстів на пропозиції (збігання виникає при ручному перекладі текстів і виражається в декількох варіантах).

Корпус паралельних текстів обслуговується також програмою ПАРТЕКС, яка має на вході два паралельні тексти – оригінал і переклад, тому вирівнювання здійснюється на рівні пропозиції. Програма видає на виході синхронізований текст, у якому послідовно за кожною пропозицією оригіналу йде переклад. Згодом програмою проводиться вирівнювання, і в результаті виходить файл Excel з паралельними пропозиціями англійською, російською й українською мовами.

Проблема колокації паралельних корпусів визначається передусім функціональним стилем тексту і стилем мови загалом. Бобкова Т. пропонує застосовувати в даному випадку семантико-синтаксичні підходи [2, с. 40]: лексико-семантичний, лексико-функціональний, лексико-граматичний і синтаксичний (Рис. 1).



Рис. 1. Семантико-синтаксичні підходи для ліквідації лексичних проблем паралельного корпусу текстів

Запропоновані підходи ідентифікації та екстракції колокацій із тексту притаманні закритим і відкритим спискам. Методика закритого списку припускає наявність заздалегідь заданих потенційних відповідей системи або наявність реєстру лексичних одиниць, у нашому випадку – стійких сполучень. Використання більш складного щодо методики напівзакритого списку ґрунтується на закритому спискові значень необхідних пошукових одиниць. Методика відкритого списку теоретично не передбачає ніяких обмежень щодо інтерпретації статистичних результатів обробки тексту. Відповідно, у традиційних лінгвістичних дослідженнях для виділення колокацій

використовується в основному методика закритого списку, а в корпусно-орієнтованих – методика відкритого списку.

Двомовний паралельний корпус текстів офіційно-ділового стилю поділяється за принципами кодування і морфологічною розміткою текстів документів з певної тематики. Складовими корпусу текстів є тематично однорідні, але різножанрові тексти: 1) офіційні документи (законодавчі акти, документи НАТО); 2) офіційно-ділові листи; 3) дискусії, що стосуються певних питань.

Лінгвістичний аналіз і переклади паралельних текстів з однієї мови на іншу вимагають уваги до морфологічної розмітки, що належать до мов різних граматичних систем. В автоматизованих системах для кожної частини мови розроблена окрема система буквено-цифрових кодів.

Створення уніфікованої системи кодів для англо-українського паралельного корпусу текстів пов'язане з об'єктивними проблемами, пов'язаними, насамперед, із різними типологічними характеристиками досліджуваних мов. Під час перекладу в українській мові граматичне значення слова виражається синтетично. Для такого способу характерне поєднання граматичних засобів вираження значення разом із словом або стійким фразеологічним сполученням.

Граматичне значення виражається за допомогою закінчення, суфікса, префікса, зміни наголосу, внутрішньої флексії. В англійських корпусах переважає аналітичний спосіб вираження граматичного значення, він проявляється за межами слова за допомогою прийменників, артиклів, допоміжних дієслів, інших службових слів, порядку слів у реченні.

Тому, головне завдання під час дослідження паралельного корпусу текстів на початковому етапі полягає в зіставленні кодів і кодових ланцюжків, щоб якісно зіставити тематично однорідні фрагменти, перекладати документи з однієї мови на іншу, створювати словники одного типу текстів; порівнювати способи вираження понять у мовній колокації різних мов.

Похибки також може створювати напівавтоматичне морфологічне кодування, за допомогою якого укладаються частотні словники словоформ і ведеться робота з вирівнювання речень оригіналу й перекладу. У створеному корпусі використовуються авторські програми для: 1) морфологічного кодування англійського дієслова на основі 15 диференційних ознак; 2) пошуку за словоформою, лемою і граматичним кодом; 3) пошуку еквівалента англійського речення в українському масиві.

Розглянемо морфологічні властивості паралельного перекладу. В процесі перекладу паралельних корпусів текстів офіційно-ділового стилю, система кодування українських прикметників більш розгалужена, ніж в англійських, зокрема фіксується рід, число, відмінок, ступені порівняння; для англійських прикметників – тільки ступені порівняння; тому в одному англійському коді з'являються кілька кодів українських іменників і прикметників. Так само розрізняються форми ступенів порівняння англійських фразеологізмів. При кодуванні українських і англійських займенників розпізнаються займенники-іменники, займенники-прикметники, а в українських текстах – їх перехід у певний розряд займенників. Не повністю висвітлюються інваріанти дієслівних форм офіційних документів. І все це пояснюється особливістю стилю.

Сьогодні корпус із морфологічною розміткою є найбільш поширеним серед інших типів корпусів. При цьому кодування охоплює не тільки ознаки частиномовної належності, а також і коди граматичних категорій, властивих для цієї частини мови. Результатом частиномовного кодування є текст, що містить однозначні морфологічні маркери. Наприклад, у процесі створення систем автоматичного аналізу російського й

українського офіційно-ділового тексту для кожної частини мови розроблено систему буквено-цифрових кодів. Тому одним із головних завдань створення двомовного паралельного корпусу текстів було зіставлення розроблених кодів щодо системи англійської мови. Натомість, створення уніфікованої системи кодів для англо-українського корпусу текстів пов'язано з об'єктивними труднощами, які пояснюються насамперед різними типологічними характеристиками досліджуваних мов.

Потенціал паралельного корпусу текстів офіційно-ділового стилю реалізовується на лексикографічному рівні.

Важливу роль відіграють гібридні підходи, як спроба подолання обмежень односторонніх підходів до вирішення завдань вилучення термінів на основі як лінгвістичних, так і статистичних елементів. Одним з основних методів оцінки ступеня термінологічності є метод автоматичного виявлення багатокomпонентних термінів у тексті. Тут використовується корпус текстів вихідної мови, на його основі формується список лексем та їх перехід у багатокomпонентні терміни. Ці терміни впорядковуються за ступенем термінологічності, який встановлюється на основі порівняння таких елементів: сумарних частот компонентів словосполучення; частоти словосполучення, включеного в довші структури; частоти головного слова і словосполучення в національному корпусі текстів. Отриманий результат оцінюється науковцем у конкретній предметній сфері. Як правило, підходи з використанням показників термінологічності засновані на об'єднанні лінгвістичної і статистичної інформації. Лінгвістична інформація складається з граматичної розмітки корпусу текстів за частинами мови, лінгвістичний фільтр обмежує тип знайдених термінів і використовує список «стоп-слів».

Тому лінгвістична база, яка реалізовує метод граматичної розмітки, містить такі компоненти: 1) інформацію про частини мови, отриману з результатів граматичної розмітки корпусу текстів; 2) власне лінгвістичний фільтр, який застосовується до розміченого корпусу текстів. До таких моделей належать невирішена комбінаторика окремих частин мови. Тоді можливе застосування як «закритого» фільтра, що дає змогу вибрати ланцюжки слів тільки конкретних типів, так і «відкритого» фільтра, у якому перераховуються лише недозволені типи ланцюжків; 3) список невпізнаних слів.

Відповідно до класифікації, статистична частина аналізує частотні характеристики ланцюжка лексичних одиниць, які упізнаються системою як терміни. Необхідність гібридного підходу визначається тим, що доступна для аналізу статистична інформація без спеціальної лінгвістичної фільтрації не є достатньою, щоб дати позитивні результати. Вибір конкретного лінгвістичного фільтра і його наповнення залежить від того, яким чином у системі збалансовані повнота і точність: перевага точності над її повнотою вимагає використання закритого фільтра, натомість перевагу повноти визначає використання відкритого. Можна зробити висновок, що методи, які використовуються в статистичних системах, варіюють точні дані, на відміну від простих підрахунків частот та зв'язку елементів колокацій, які можуть бути в структурі лексеми, що визначається як термін.

**Висновки.** Розглянувши похибки та позитивні показники паралельного корпусу текстів офіційно-ділового стилю, можемо стверджувати, що його побудова здійснюється для конкретної мови, спеціальних цілей, встановлення комплексу лінгвістичних і статистичних параметрів. Головні завдання, які необхідно виконати – знайти похибки в кодовій структурі термінів та ліквідувати нерівномірність колокації.

## Література

1. Бобкова Т., Перебийніс В., Сорокін В. Частотні словники паралельних текстів / Т. Бобкова, В. Перебийніс, В. Сорокін // Людина. Комп'ютер. Комунікація: [зб. наук. праць]. – Львів: Вид. Національного університету «Львівська політехніка», 2008. – С. 158–160.
2. Бобкова Т. В. Лексикографический корпус как источник для словарей нового типа. / Т. В. Бобкова // In A. Diomidova, L. Kamičaitytė, J. Radavičiūtė, ed. Žmogus kalbos erdvėje, no. 8, 2015. – P. 15–26.
3. Бобкова Т. В. Теоретико-методологічні підходи до вивчення колокацій у сучасному мовознавстві / Т. В. Бобкова // Вісник Київського національного лінгвістичного університету. Сер.: Філологія. – 2014. – т. 17, № 2. – С. 14–22.
4. Большакова Е. И. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие / Е. И. Большакова, Э. С. Клышинский, Д. В. Ландэ. А. А. Носков, О. В. Пескова, Е. В. Ягунова. – М.: МИЭМ, 2011. – 272 с.
5. Борисова Е. Г. Фразеологическое значение в устойчивых словосочетаниях / Е. Г. Борисова, О. В. Захарова // Филологические науки, №4. 1994. – С. 77–84.
6. Добровольский Д. О. Корпус параллельных текстов: архитектура и возможности использования / Д. О. Добровольский, А. А. Кретов, С. А. Шаров // Национальный корпус русского языка: 2003-2005. – М.: Индрик, 2005. – С. 263–296.
7. Касьяненко М. К. Принципи побудови корпусу текстів різних функціональних стилів / М. К. Касьяненко, К. М. Лебедев, П. М. Петренко // Науковий вісник Волинського національного університету імені Лесі Українки. Вежа: 2009, – №12. – С. 25-28.
8. Клименко С. В. Логические индукция и дедукция как принципы отражения предметной области в корпусе текстов [Электронный ресурс] / С. В. Клименко, В. В. Рыков, 2001. – [www.dialog-21.ru/materials/archive.asp?id=6915&y=2001&vol=6078](http://www.dialog-21.ru/materials/archive.asp?id=6915&y=2001&vol=6078).
9. Ягунова Е. В., Пивоварова Л. М. От коллокаций к конструкциям. Acta linguistica petropolitana. Труды Института лингвистических исследований. РАН. [Электронный ресурс] available at: [http://www.webground.su/data/lit/pivovarova\\_yagunova/Ot\\_kollokatsiy\\_k\\_konstruktsiyam.pdf](http://www.webground.su/data/lit/pivovarova_yagunova/Ot_kollokatsiy_k_konstruktsiyam.pdf) [Accessed April 2015].
10. Baker M. Corpora in Translation Studies. An Overview and Suggestions for Future Research / M. Baker // Target. – 1995. – Vol. 7:2. – pp. 223–243.
11. Delpech E., Daille B. Dealing with lexicon acquired from comparable corpora: validation and exchange // Proceedings, 9th Conference on Terminology and Knowledge Engineering (TKE). Fiontar, Dublin City University, 2010. – P. 229-223.
12. Granger S. The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies // S. Granger, J. Lerot and S. Petch-Tyson (eds). Corpus-based Approaches to Contrastive Linguistics and Translation Studies. – Amsterdam: Rodopi, 2003. – pp. 17–29.
13. Sinclair J. Corpus and Text – Basic Principles [Text] / J. Sinclair // Developing Linguistic Corpora: a Guide to Good Practice, ed. M. Wynne. – Oxford: Oxbow Books. – 2005. – pp. 1-16.
14. Véronis J. From the Rosetta Stone to the Information Society: A Survey of Parallel Text Processing // J. Véronis (ed.). Parallel Text Processing: Alignment and Use of Translation Corpora. Kluwer Academic Publishers, 2000. – pp. 1–24.

## References

6. Bobkova, T., Perebyinis, V. and Sorokin, V. (2008), *Frequency dictionaries of parallel texts* [Chastotni slovnyky paralelnykh tekstiv], Liudyna. Kompiuter. Komunikačija: [zb. nauk. prats], Lviv, pp. 158–160.
7. Bobkova, T. V., (2015), *Lexicographic corpus as a source for new type of dictionaries* [Leksikograficheskiy korpus kak istochnik dlya slovarey novogo tipa]. In A. Diomidova, L. Kamičaitytė, J. Radavičiūtė, ed. *Žmogus kalbos erdvėje*, No. 8, pp. 15–26.
8. Bobkova, T. V., (2014), *Theoretical and methodological approaches to the study of collocation in modern linguistics* [Teoretyko-metodolohichni pidkhody do vyvchennia kolokatsii u suchasnomu movoznavstvi], *Visnyk Kyivskoho natsionalnoho linhvistychnoho universytetu*. Ser.: Filolohiia, No. 2, pp. 14–22.
9. Bolshakova, E. Y., Klyshinskiy, E. S., Lande., D. V., Noskov, A. A. and Peskova, O. V., (2011), *Automatic processing of texts in natural language and computer linguistics: a tutorial* [Avtomaticheskaya obrabotka tekstov na estestvennom yazyke i komp'yuternaya lingvistika: uchebnoe posobie], Moscow: MIEM, 272 p.
10. Borisova, E. G., (1994), *Phraseological meaning in collocations* [Frazеologicheskoe znachenie v ustoychivyykh slovosochetaniyakh], *Filologicheskie nauki*, No. 4, pp. 77–84.
11. Dobrovol'skiy, D. O., Kretov, A. A. and Sharov, S. A., (2005), *Corpus of parallel texts: architecture and usage possibilities* [Korpus parallel'nykh tekstov: arhitektura i vozmozhnosti ispol'zovaniya], *Nacional'nyj korpus russkogo jazyka: 2003–2005*, Moscow: Indrik, pp. 263–296.
12. Kasianenko, M. K., Lebediev, K. M. and Petrenko, P. M., (2009), *Principles of construction the corpus texts of different functional styles* [Pryntsypy pobudovy korpusu tekstiv riznykh funktsionalnykh styliv], *Naukovyi visnyk Volynskoho natsionalnoho universytetu imeni Lesi Ukrainky*. No. 12, pp. 25–28.
13. Klimenko, S. V. and Rykov, V. V., (2001), *Logical induction and deduction as the principles reflected in the domain corpus* [Logicheskie indukciya i dedukciya kak principy otrazheniya predmetnoj oblasti v korpuse tekstov]. [online] available at: <<http://www.dialog-21.ru/materials/archive.asp?id=6915&y=2001&vol=6078>>.
14. Jagunova, E. V. and Pivovarova, L. M., (n.d.) *From collocations to structures* [Ot kollokacij k konstrukcijam]. [online] *Acta linguistica petropolitana. Trudy Instituta lingvisticheskikh issledovanij: RAN*. available at: <[http://www.webground.su/data/lit/pivovarova\\_yagunova/Ot\\_kollokatsiy\\_k\\_konstruktsiyam.pdf](http://www.webground.su/data/lit/pivovarova_yagunova/Ot_kollokatsiy_k_konstruktsiyam.pdf)> [Accessed April 2015].
15. Baker, M., (1995), *Corpora in Translation Studies. An Overview and Suggestions for Future Research*. *Target: International journal of translation studies*, Vol. 7:2, pp. 223–243.
16. Delpech, E. and Daille, B., (2010), *Dealing with lexicon acquired from comparable corpora: validation and exchange*. *Proceedings, 9th Conference on Terminology and Knowledge Engineering (TKE)*. Fiontar, Dublin City University, 2010, pp. 229–223
17. Granger, S., (2003), *The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies*. In S. Granger, J. Lerot and S. Petch-Tyson (eds). *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam: Rodopi, pp. 17–29.
18. Sinclair, J., (2005), *Corpus and Text – Basic Principles*. In ed. M. Wynne *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxbow Books, pp. 1–16.

19. Véronis, J., (2000), *From the Rosetta Stone to the Information Society: A Survey of Parallel Text Processing*. In ed. J. Véronis. *Parallel Text Processing: Alignment and Use of Translation Corpora*. Kluwer Academic Publishers, pp. 1–24.

**Ю. И. Демьянчук,**

*Львовский государственный университет безопасности жизнедеятельности,  
кафедра технического перевода*

### **ПАРАЛЛЕЛЬНЫЙ КОРПУС ТЕКСТОВ: ТЕОРЕТИКО-МЕТОДОЛОГИЧЕСКИЙ И ЛЕКСИКОГРАФИЧЕСКИЙ АНАЛИЗ, ПРИНЦИПЫ ПРИМЕНЕНИЯ**

В статье рассматриваются теоретико-методологические аспекты обработки двух или более текстов на основе корпусного перевода. Указывается на положительный лексикографический уровень параллельного корпусного перевода (обогащение текста устойчивыми словосочетаниями, идиомами, уточнениями, сроками, терминологическими выражениями и т.д.). Сделан вывод о сохранении в отдельных конструкциях-коллокациях устойчивых лексических единиц, поддаются автоматизации и графической коррекции.

**Ключевые слова:** коллокация, параллельный корпусный подход, лексикографический анализ, автоматизация текста, авторский перевод.

**Yu. I. Demianchuk,**

*Lviv State University of Life Safety,  
Technical Translation Department*

### **PARALLEL CORPUS OF TEXTS: THEORETICAL, METHODOLOGICAL AND LEXICOGRAPHICAL ANALYSIS OF PRINCIPLES**

The article deals with theoretical and methodological aspects of processing two or more texts from the corpus translation, indicates a positive lexicographical level of parallel corpus translation (text enrichment stable phrases, idioms, specification, terms, terminology expressions, etc. (rich text set phrases, idioms, updates, deadlines, terminology and expressions etc., outputs to maintain in the individual designs collocation stable lexical units, which are amenable to automation and graphical correction).

The main aim of the article is to identify the theoretical and methodological principles on the use of parallel corpus of texts, explain lexicographical changes of individual units of the collocation of text.

Conceptual objectives are: to prove the relationship between corpus approach and parallel corpus of text; to determine the principles of construction the parallel corpus of texts of different functional styles; to determine errors of parallel corpus approach (align, differences, saving recognized text, lexical, grammatical and phraseology differences).

The main task in the study of parallel corpus of texts at an early stage is to compare the codes and code strings to qualitatively compare thematically homogeneous fragments, translate documents from one language to another, creating a type of dictionary texts; compare ways of expressing concepts in language collocation of different languages.

Having examined the errors and the positive performance of parallel corpus of texts of the official-business style, we can assert that its construction is carried out for a specific language, special purposes, establish complex linguistic and statistical parameters. The main



tasks that must be done are to find errors in the code structure terms and to eliminate uneven collocation.

**Key words:** collocation, parallel corpus approach, lexicographical analysis, automation text, author's translation.

УДК 811. 161. 2'42

**Н. Я. Лужецька,**

*Дрогобицький державний педагогічний університет*

*імені Івана Франка,*

*аспірант кафедри філологічних дисциплін*

*та методики їх викладання*

## **ПОЛІТИЧНА МОВА У ЛІНГВОПОЛІТОЛОГІЧНІЙ ХАРАКТЕРИСТИЦІ ІВАНА ФРАНКА**

Фундаментальний підхід до характеристики політичної мови спостерігаємо у політичному дискурсі Івана Франка – засновника української „школи політичного мислення”. Мета наукової розвідки – через лінгвополітологічні погляди мислителя висвітлити особливості тогочасної політичної мови.

**Ключові слова:** політичне мислення, політична мова, комунікація, Іван Франко, лінгвополітологія, маніпуляція свідомістю.

*Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями.* Дослідження лінгвополітологічної проблематики зумовлене концептуальним спрямуванням сучасної лінгвістичної науки до антропоцентричної парадигми, що базується на тріаді „культура – свідомість – мова”. Оскільки „ми можемо дістатися до думки тільки через слова” [1, с. 293], то метою лінгвополітології закономірно постає вивчення за мовленнєвою поведінкою специфіки політичного мислення. У соціально-психологічних характеристиках політичних комунікацій і комунікантів гострої актуальності набуває проблема використання мови як засобу маніпуляції свідомістю.

*Аналіз останніх досліджень і публікацій.* Реалізація маніпулятивних стратегій і тактик особливо поширена у політичній комунікації. Окреслена проблема стала об'єктом досліджень американських і західноєвропейських лінгвістів (П. Серіо, Р. Мідоу, Д. Болінджер, Р. Блакар, Х. Вайнрих, Р. Водак, Т. ван Дейк, Р. Лакофф, Л. Лассвел, В. Клемперер та ін.). У 80-ті рр. ХХ ст. У 90-ті рр. до проблематики політичної лінгвістики звернулися у колишніх країнах СРСР: спершу досліджувалася тоталітарна мова радянської епохи, що характеризувалася гаслами, ритуалізованістю, імітацією свята (А. Баранов, Ю. Левін, Н. Купина, С. Єрмоленко, Є. Рабинович, В. Алпатов та ін.), згодом вектор студій було спрямовано на теоретичне розроблення мовних особливостей організації політичного дискурсу (Е. Лассан, О. Шейгал, А. Белова, О. Романов, В. Козловський, О. Паршина, Л. Синельникова та ін.). Український політичний дискурс останніх років став предметом уваги не лише вітчизняних лінгвістів (Г. Яворської, Л. Ставицької, С. Єрмоленко, Н. Кондратенко та ін.), а й представників української діаспори в Канаді та США (Б. Гарасиміва, Л. Павлюк, В. Полковського й ін.) [4, с. 565].