

полная фонологическая ассимиляция, то есть они отражают особенности произношения языка-источника. Возникает необходимость осуществить анализ фонетических структур поздних французских заимствований в аспекте британской произносительной нормы в сопоставительном плане с американским произносительным вариантом.

**Ключевые слова:** фонемная структура, акцентная структура, свободное фонемное и акцентное варьирование, фонологическая ассимиляция заимствований.

**Komar Liudmyla. Peculiarities of Phonemic and Accentual Structures of French Late Borrowings in British Pronouncing Norm.** The paper describes the peculiar features of the phonetic structure of late French borrowings in British pronunciation norm. The relevance of comprehensive study of French words in English vocabulary is due to the large number of such words in real English usage, lack of knowledge about the specific phonological standartization. The free phonemic and accentual variation, availability of foreign phonemes, phonotactic specifications determine peculiarities of phonetic structure of French words. In general, the subsystem of French borrowings in Modern English is characterized by a partial, rather than full phonological assimilation, that is they reflect the peculiarities of pronunciation of source language. There is a need to analyze phonetic structures of late French borrowings in British pronunciation norm in comparison with American pronunciation variant.

**Key words:** phonemic structure, accentual structure, free phonemic and accentual variation, phonological assimilation of borrowings.

Стаття надійшла до редколегії  
13.02.2013 р.

УДК 81'322.2 : 81'33

**Ігор Кульчицький,  
Ігор Ліхнякевич,  
Юлія Данчевська**

### **Деякі аспекти створення та використання паралельних корпусів**

У статті розглянуто основні аспекти створення та функціонування паралельних корпусів. Із розвитком комп'ютерних технологій зріс і попит та можливість для вдосконалення корпусних методик, зокрема і стосовно паралельного корпусу. Автори здійснили огляд відмінностей термінологічного означення поняття «паралельний корпус» у західній науці. Окрім цього, окреслено основні вимоги до створення паралельного корпусу, а також ті труднощі, які можуть виникнути при цьому в розробників. Проаналізовано основні здобутки та сфери застосування цього типу корпусу для різноманітних досліджень у галузі контрастивної лінгвістики, перекладознавства, а також лексикографії.

**Ключові слова:** паралельний корпус, вирівнювання, програмне забезпечення, еквівалент перекладу, контекст, словоформа, розмітка, двомовні словники.

**Постановка наукової проблеми та її значення.** Міжнародна комунікація та процеси глобалізації спричинили бурхливий розвиток перекладознавчих та контрастивних студій. Корпусна лінгвістика та багатомовні корпуси мовних даних зокрема посідають у цьому важливе місце. Завдяки таким своїм якостям, як доступність і легкість у редагуванні та доповненні, вони відкривають практично невичерпні можливості для різноманітних розвідок у галузі науки про мову, починаючи від отримання квантитативних даних та створення дидактичних матеріалів і до створення складних лексикографічних систем та баз даних, удосконалення систем опрацювання природної мови та машинного перекладу. Не в останню чергу такий поступ став можливим завдяки появі різних типів корпусів, що набуло великого розвитку з кінця минулого століття. Відповідно значна кількість різнопланових корпусів спонукала науковців до створення їх типології за певними ключовими параметрами (А. Баранов, Л. Ричкова та ін.). Зокрема, виокремлюють такі категорії: за метою та призначенням; за типом текстового матеріалу; за типом програмного оброблення [5, с. 83–101]

**Аналіз останніх досліджень цієї проблеми.** Згідно з вищезазначеною класифікацією паралельний корпус, який є **об'єктом** нашого дослідження, виокремлюють за типом текстового матеріалу,

однак існують й інші думки із цього питання, які висловили західні вчені (В. Тойберт, М. МакКарті, Е. МакЕнері й ін.). Окрім цього, паралельний корпус є важливим засобом для вдосконалення перекладацьких навичок, створення двомовних словників, досліджень у контрастивній лінгвістиці тощо.

**Мета статті** – визначення місця паралельного корпусу в корпусній таксономії та його прикладне значення у вивченні мовних явищ.

**Виклад основного матеріалу й обґрунтування отриманих результатів дослідження.** Паралельні корпуси належать до дво- або багатомовних корпусів, залежно від кількості мов перекладу вихідних текстів. Західні вчені висловлюють суперечливі думки щодо того, що називати терміном «паралельний корпус». Зокрема, С. Грейнджер у статті «Comparable and translation corpora in cross-linguistic research. Design, analysis and applications», аналізуючи використання терміна «паралельний корпус» різними вченими, зазначає, що в контрастивній лінгвістиці виокремлюють два основних типи корпусів: 1) корпуси оригінальних текстів та їх переклад однією чи кількома мовами, які називають перекладацькими; 2) корпуси, які містять оригінальні тексти кількома мовами, відібрані за такими критеріями як час створення, адресат тощо, який ще називають порівняльним корпусом. Тобто іноді паралельним корпусом називають перекладацький, порівняльний та порівняльно-перекладацький корпуси. Перекладознавці виокремлюють перекладацький, паралельний та порівняльний корпуси. Порівняльним корпусом вони називають два окремі корпуси текстів, де один корпус – це тексти досліджуваної мови, а інший – переклади цією мовою текстів із заданої мови-оригіналу, а перекладацьким – корпус перекладених текстів. Натомість у контрастивній лінгвістиці порівняльним корпусом називають зіставлені оригінальні тексти різними мовами, а перекладацьким – одномовні корпуси оригінальних та перекладених текстів однією мовою. У рамках перекладознавчих студій паралельним корпусом називають корпус вихідних текстів, вирівняних із їх відповідними перекладами. Це те визначення, під яким у контрастивній лінгвістиці розуміють перекладацький корпус [11, с. 15]. Британський дослідник Е. МакЕнері також зазначає термінологічну неусталеність паралельного корпусу, який він відносить до багатомовних корпусів. Він виокремлює такі його типи: А) вихідні тексти та їх переклади (проекти *CRATER*, *Canadian Hansard*, *ParaSol*); Б) одномовний підкорпус, який містить тексти, дібрані за одним і тим самим принципом (проект *The Aarhus corpus of contract law*); В) поєднання двох вищезазначених типів (проекти *ENPC*, *EMILE*). Учений зазначає, що, у першу чергу, варто визначити за яким критерієм будемо оцінювати корпус – за кількістю мов (одно-, дво- чи багатомовний) чи за наповненням корпусу (перекладний чи порівняльний) – та дотримуватися цього критерію при аналізі будь-якого паралельного корпусу [12, с. 1–2]. Отже, проаналізувавши кілька точок зору, ми використовуємо термін «паралельний корпус» у його такому визначенні – це підмножина текстів мовою-оригіналу та одна чи кілька підмножин текстів перекладу мови-оригіналу на цільову/і мову/и [2, с. 57–60; 5, с. 87]. Серед паралельних корпусів виокремлюють двомовні та багатомовні, а також односпрямовані (наприклад, або з української на англійську, або з англійської на українську), двоспрямовані (містять і оригінальні тексти українською, перекладені англійською, і англійські оригінальні тексти, перекладені українською; наприклад, англо-португальський проект *COMPARA* - <http://www.linguateca.pt/COMPARA/>) та багатоспрямовані (один і той самий текст українською мовою, представлений англійською, німецькою, французькою та навпаки). Прототипи паралельного корпусу, зокрема корпус слухань парламенту ЄС (*The Europarl Corpus* – <http://www.statmt.org/europarl/>), власне належить до третього типу, оскільки одночасно містить тексти кількома різними мовами, які відповідають один одному.

Для того, щоб паралельний корпус став корисним зняряддям, він повинен бути грамотно створений та забезпечений відповідним програмним інструментарієм. Спочатку корпуси паралельних текстів просто містили тексти, переважно нормативного характеру, перекладені різними мовами (вищезгаданий *The Europarl Corpus*) і цього було достатньо, адже існували вони для ознайомлення з певними документами, які мають строгі умови перекладу. Натомість із розвитком технологій та появою паралельних текстів художньої літератури зросли і складність запитів до паралельного корпусу та, відповідно, їх структура та функціональність. Основною вимогою до паралельного корпусу є вирівнювання тексту-оригіналу з текстом-перекладом. Власне, це і є основною складністю при створенні паралельного корпусу, адже зрозуміло, що далеко не кожне слово має свій точний еквівалент у іншій мові. Тому, зазвичай, вирівнювання здійснюють на рівні абзаців/параграфів та речень, для чого створюють спеціальне програмне забезпечення. Хоча і тут можливі певні невідповідності,

адже в різних мовах та в різних виданнях іноді використовують різні способи графічного оформлення тексту (приміром, різні способи передачі переходу від прямого мовлення персонажів до авторського мовлення), тому програми зазвичай повідомляють про такі випадки. Детальніший огляд роботи програмного забезпечення – тема наступних наших розвідок.

Для пошуку конкретних словоформ у паралельному корпусі не потрібно спеціальної розмітки, адже завдання зі сфери лексичної семантики можна розв'язувати і за допомогою корпусів, які ще не пройшли етап морфологічного та синтаксичного анотування. Про це свідчать дослідження Е. Мак-Енері, Д. Добровольського й ін. [7; 12].

Як слушно зауважує В. Тойберт [13, с. 184], «будь-хто, якщо коли-небудь перекладав якийсь текст рідною чи іноземною мовою, розуміє, що ми перекладаємо не слово в слово» і водночас зазначає, що двомовні перекладні словники вже не є універсальними помічниками як раніше. Із цим погоджуються і російські вчені М. Добровольський [6] та Є. Сосніна [10], котрі виявляють такі важливі відмінності: двомовні словники – це набори лексичних еквівалентів (звичайні словники) чи термінів (спеціалізовані словники та термінологічні), представлені творцями словників як варіанти перекладу. Паралельні корпуси є свого роду збірками стратегій та еквівалентів перекладу, якими послуговувались перекладачі. Обираючи еквівалент перекладу зі звичайного двомовного словника, перекладач повинен оцінити, чи підходить запропонований варіант до нового контексту, перевіряючи доречність кожного із запропонованих. Паралельний корпус пропонує приклади способів перекладу схожих проблематичних моментів, представляючи їх у «дзеркальному» відображенні та у відповідному контексті, що значно спрощує вибір найкращого варіанта. Паралельний корпус наочно демонструє як саме перекладач впорався з такими проблемами. Завдяки цьому паралельні корпуси можуть забезпечити нас інформацією, яку двомовні словники зазвичай не містять, адже вони пропонують також і переклад безеквівалентної лексики [1; 3; 4].

Єдиним недоліком такого підходу є те, що не всі переклади є якісними, і паралельний корпус може вказати і на такі явища, як неадекватність перекладу. Хоча загалом найпоширенішими наразі є паралельні корпуси власне класичних творів художньої літератури, адже їм характерний високий літературний рівень і якісні переклади кількома мовами. Серед таких відомими є міжнародні проекти ParaSol (режим доступу – <http://parasol.unibe.ch/>) та InterCorp (режим доступу – <http://www.korpus.cz/intercorp/?req=doc:uvod>), основну частку яких становлять художні тексти, англо-чеський корпус паралельних текстів Kacenska (режим доступу – <http://www.phil.muni.cz/angl/kacenska/kachna.html>), КоПарТ (частина корпусу російської мови), російсько-болгарський паралельний корпус (96,7 % текстів художнього стилю; режим доступу – [http://rbcorpus.com/structura\\_rus.php](http://rbcorpus.com/structura_rus.php)), російсько-німецький корпус роману Ф. М. Достоєвського «Ідіот» як частина корпусу Австрійської академії наук (режим доступу – [http://www.oeaw.ac.at/fackellex/geplantes/english/fut2c\\_e.html](http://www.oeaw.ac.at/fackellex/geplantes/english/fut2c_e.html)), російсько-фінський корпус художніх творів «ПарРус» [Михайлов], паралельний корпус перекладів «Слово о полку Ігоревім» (режим доступу – <http://nevmenandr.net/slovo/>), паралельний корпус російських літописів (різних редакцій) (режим доступу – [http://manuscripts.ru/mns/cred.analyzer?koll=62133570&f\\_type=14001](http://manuscripts.ru/mns/cred.analyzer?koll=62133570&f_type=14001)) та ін.

Як зазначає М. Михайлов [8, с. 32–38], для створення паралельного корпусу художніх творів характерно кілька нюансів: на відміну від загального корпусу, корпус художніх творів повинен містити повні тексти; до корпусу текстів повинні ввійти твори відомих та знакових письменників, які внесли значний вклад у певну культуру та її мову; твори можуть бути створені за значний проміжок часу.

Пошук вихідного тексту може стати проблематичним, адже ідеальним було б включення в корпус того видання, за яким здійснювався переклад, що, у свою чергу, може виявитись складним завданням, якщо твір було написано доволі давно й він витримав кілька перевидань та редагувань чи рецензувань. Водночас іноді дотримання авторського права не дає змоги широко застосовувати ці корпуси в освітніх та дослідницьких цілях. Цим фактом можна пояснити появу невеликих корпусів, які містять твори певного письменника чи певного періоду. Проте в поєднанні з одномовними загальними корпусами, вони можуть стати надзвичайно корисними, адже вони міститимуть специфічну лексику, граматичні конструкції, які відсутні в загальномовних корпусах. Унікальність корпусу художніх текстів полягає і в тому, що вони характеризують слово в контексті, на відміну від вузько-спеціалізованої термінологічної лексики, якій властива нормативна фіксація єдиного прийнятного еквіваленту [9]. Окрім цього, при лінгвістичному аналізі художнього тексту великої уваги надають

особливостям ідіостилю письменника. Індивідуальність мови може виявлятися навіть на рівні вибору відмінкової форми іменника. Проте для того щоб виявити специфіку художнього тексту, потрібно знати фон, на якому вона виявляється. При лінгвістичному аналізі художнього тексту можна використати корпус як фон для вивчення ідіостилю, запропонувавши схарактеризувати як письменник створює індивідуальну картину світу за допомогою певних граматичних форм ключових слів.

Ще одне, не менш важливе, прикладне значення мають паралельні корпуси при створенні словників різного складу та призначення. Завдяки значному поступу в розвитку комп'ютерних технологій, друге дихання та новий поштовх у розвитку отримала і лексикографія. Будучи прикладною дисципліною, основним завданням якої є розроблення методів створення словників, лексикографія значно розширила свої можливості завдяки спеціальним програмам, базам даних, програмам обробки тексту, які дають змогу в автоматичному режимі формувати словникові статті, зберігати словникову інформацію та обробляти її. Різні комп'ютерні лексикографічні програми можна розділити на дві великі групи: програми підтримки лексикографічних робіт і автоматичні словники (АС) різних типів, що включають лексикографічні бази даних. Власне питанням комп'ютерної лексикографії докладно та розширено досліджує В. Широков (див. монографію «Елементи лексикографії», 2005), а також Український мовно-інформаційний фонд НАН України під його керівництвом. Використання корпусу паралельних текстів у двомовній лексикографії дає змогу не тільки максимально автоматизувати відбір словосполучень, але також служить для:

- збагачення набору словникових статей за рахунок вибору вільних словосполучень, використовуваних у вихідних текстах, що надзвичайно важливо для тих, хто перекладає на мову, яка не є рідною;
- уточнення використання конкретних словосполучень у текстах певної предметної галузі;
- верифікації значень лексичних одиниць, зафіксованих у двомовних словниках, особливо в тому, що стосується ідіом та термінологічних виразів;
- виділення стійких словосполучень та ідіом, які доцільно ввести в словник конкретної галузі знань.

**Висновки.** Паралельний корпус характеризується складною композицією та програмним забезпеченням, проте в разі грамотного створення може слугувати багатим джерелом для різноманітних лінгвістичних студій і для спеціалістів-філологів та перекладачів, і для навчання студентів. Окрім того, паралельний корпус може слугувати джерелом для вдосконалення лексикографічних праць і систем машинного перекладу.

**Перспективи подальшого дослідження.** У цій статті ми описали основи паралельного корпусу, починаючи від термінологічних аспектів та закінчуючи прикладним значенням паралельного корпусу. Перспективними, на нашу думку, будуть детальніші дослідження чинного програмного забезпечення для створення паралельних корпусів, зокрема художніх творів певного автора, а також аналіз уже чинних проектів, які можуть бути зразком для створення аналогічних проектів в Україні.

#### *Джерела та література*

1. Андреева Е. Г. Анализ переводческих соответствий на материале параллельного корпуса текстов // Компьютерная лингвистика и интеллектуальные технологии : тр. Междунар. конф. «Диалог 2006». – М. : РГГУ, 2006. – С. 26–30.
2. Баранов А. Н. Введение в прикладную лингвистику : [учеб. пособие] / А. Н. Баранов. – [2-е изд., испр.]. – М. : Едиториал УРСС, 2009. – 360 с.
3. Башкова И. В. Данные корпуса на занятиях по синтаксису и лингвистическому анализу художественного текста / И. В. Башкова // Национальный корпус русского языка и проблемы гуманитарного образования. – М. : [б. и.], 2007. – С. 207–220.
4. Беляева Л. Н. Корпус параллельных текстов как база данных прикладной лексикографии / Л. Н. Беляева // Корпусная лингвистика 2008 : тр. Междунар. конф. – СПб. : Петербург. гос. ун-т, ф-т филологии и искусств, 2008. – с. 31–40.
5. Демська О. М. Текстовий корпус: ідея іншої форми / О. М. Демська. – К. : Вид. дім «Києво-Могилянська академія», 2011. – 284 с.
6. Добровольський Д. О. Корпус параллельных текстов и литературный перевод / Д. О. Добровольский // НТИ. Сер. 2, 10. – 2003. – С. 13–18.

7. 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. Результаты и перспективы. – М. : Индрик, 2005. – С. 263–296.
8. Михайлов М. Н. Параллельные корпуса художественных текстов: принципы составления и возможности применения в лингвистических и переводоведческих исследованиях / М. Н. Михайлов. – Тампере : Тампер. ун-т, 2003. – 348 с.
9. Потемкин С. Б. Проблемы разработки параллельного корпуса переводов русской классики / С. Б. Потемкин // Вестн. Воен. ун-та. – Вып. 2. – 2012. – С. 138–145.
10. Соснина Е. П. Параллельные корпуса в обучении языку и переводу [Электронный ресурс] / Е. П. Соснина. – Режим доступа : [http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus\\_education\\_translation/](http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_education_translation/)
11. Granger S. Comparable and translation corpora in cross-linguistic research. Design, analysis and applications / S. Granger // Journal of Shanghai Jiaotong University. – 2010. – Vol. 2. – P. 14–21.
12. McEnery A. Parallel and comparable corpora: What are they up to? 2007 // Incorporating Corpora: Translation and the Linguist / A. McEnery, Z. Xiao ; G. James, G. Anderman (eds.). – Clevedon, UK : Multilingual Matters, (Translating Europe), 2007. – P. 131–145.
13. Teubert W. Units of Meaning, Parallel Corpora, and their Implications for Language Teaching / W. Teubert // Language and Computers, Applied Corpus Linguistics. A Multidimensional Perspective. – 2004. – P. 171–189.

**Кульчицкий Игорь, Лихнякевич Игорь, Данчевская Юлия. Некоторые аспекты составления и использования параллельных корпусов.** В статье авторы рассматривают основные аспекты построения и функционирования параллельного корпуса. Вместе с развитием компьютерных технологий возросли спрос, а также возможности совершенствования корпусных методик, в частности и относительно параллельных корпусов. Авторами осуществлен обзор отличий терминологического значения понятия «параллельный корпус» в западной науке. Кроме того определены основные требования к составлению корпуса, а также те трудности, которые могут возникнуть при этом у разработчиков. Авторами также проанализированы основные достижения и сферы применения данного типа корпуса для разных исследований в области контрастивной лингвистики, переводоведения, в частности и в лексикографии.

**Ключевые слова:** параллельный корпус, выравнивание, программное обеспечение, эквивалент перевода, контекст, словоформа, разметка, двуязычные словари.

**Kulchytskyj Ihor, Likhniakievych Ihor, Danchevska Yuliya. Some Aspects of Parallel Corpora Creation and Use.** In the article the authors discuss the main aspects of parallel corpora creation and functioning. Along with the development of computer technologies there is an increasing demand for as well as the ways for improving corpus methodologies, particularly those concerning parallel corpus. The authors have conducted the review of terminological issues of the notion «parallel corpus» in western scientific thought. Moreover the main requirements and specifications while creating a parallel corpus have been outlined as well as possible difficulties that may occur during its compilation. The authors have analyzed major achievements and practical application of such corpus type for various researches in the field of contrastive linguistics, translation studies, as well as in lexicography.

**Key words:** parallel corpus, alignment, translation equivalent, context, word form, tagging, bilingual dictionary.

Стаття надійшла до редколегії  
04.03.2013 р.

УДК 811.111:81'374

Анастасія Шевчук

### Лексикографічна репрезентація гіперонімо-гіпонімічних зв'язків (на матеріалі дефініцій зоонімів)

Статтю присвячено проблемі вираження гіперонімо-гіпонімічних зв'язків у дефініціях зоонімів, вибраних із трьох тлумачних словників сучасної англійської мови. Основним способом репрезентації цих зв'язків є родовидовий тип тлумачення через найближчий гіперонім із вказівкою на видові відмінності. Показано, що в лексикографічній практиці загальної схеми тлумачення лексем-зоонімів часто не дотримуються і тому вона потребує уніфікації.

**Ключові слова:** гіперонім, гіпонім, зоонім, дефініція, лексикографія.

©Шевчук А., 2013