

УДК 81'322

*Олена Сірук*  
*Київський національний університет імені Тараса Шевченка,*  
*Іван Держанський*  
*Інститут математики й інформатики*  
*Болгарської академії наук*

## **ЛЕКСИЧНІ ПЕРЕКЛАДНІ ЕКВІВАЛЕНТИ В БОЛГАРСЬКИХ І УКРАЇНСЬКИХ ПАРАЛЕЛЬНИХ ТЕКСТАХ**

*Мова йде про укладання Корпусу болгарських і українських паралельних текстів та його застосування для дослідження лексичної семантики українських і болгарських перекладних еквівалентів.*

**Ключові слова:** корпусна лінгвістика, паралельні тексти, корпус, лексична семантика, перекладні еквіваленти.

### **Паралельні корпуси в лексичних дослідженнях**

На сучасному етапі розвитку лінгвістики для зіставного аналізу лексичної семантики, як і загалом для міжмовних досліджень, усе частіше застосовують корпуси паралельних – оригінальних і перекладних – текстів [Cysouw, Wälchli]. Такі корпуси розробляються для багатьох мов, зокрема слов'янських (наприклад, двомовні корпуси з українськими текстами в складі Національного корпусу російської мови <http://ruscorpora.ru>, Корпус паралельних російських і болгарських текстів <http://rbcorpus.com> тощо). Однак українська та болгарська мови досі не ставали об'єктом зіставного корпусного аналізу, а тексти цими мовами не були об'єднані в паралельний корпус. Автори статті мають на меті розповісти про свій досвід створення такого корпусу, а також про деякі аспекти його застосування для лексико-семантичного та статистичного аналізу перекладних еквівалентів.

### **Корпус болгарських і українських паралельних текстів (КУБ)**

Двомовний корпус містить болгарські й українські паралельні тексти, наявні в електронних бібліотеках або відскановані та розпізнані нами з паперових видань. Цим зумовлюється перевага в корпусі творів художньої літератури, зокрема романів, які домінують у таких джерелах.

Оскільки оригінальних та перекладених паралельних текстів для української та болгарської мов наразі обмаль, а тим більше в комп'ютерному вигляді та мережевому доступі, було вирішено використовувати як матеріал для корпусу також паралельні болгарські й українські художні переклади з інших мов. Тобто поняття паралельності застосовується нами в ширшому значенні, аніж його тлумачать, приміром, розробники паралельних корпусів НКРЯ [Добровольский, Кретов, Шаров], російсько-болгарського, російсько-словацького корпусу [Гарабик, Захаров], лексикографи [Лендау : 314], залучаючи до паралельних корпусів тільки пари текстів “оригінал – переклад”. Інші дослідники ототожнюють термін “паралельні тексти” із терміном “бітекст”, яким позначають просто дві версії одного тексту, зазвичай різними мовами [Vitas, Krstev, Laporte]. Ми розуміємо під паралельним корпусом як “єдність підмножини оригінальних текстів та підмножини їх перекладів на іншу(і) мову(и)” [Демська : 90], так і те, що дослідники називають взаємопаралельним корпусом. Такі корпуси “містять як оригінали, так і переклади на конститутивні мови корпусу” [Демська : 100]. Тож наш корпус має декілька секторів, кожен з яких охоплює паралельні болгарські й українські тексти, перекладені з однієї мови. Усі сектори приблизно однакові за розміром. На сьогодні українська частина містить приблизно по 700 тис. слововживань із кожного сектору, болгарська – відсотків на 15 більше, з огляду на граматичні особливості мови та стилістичні вподобання багатьох перекладачів (це можна побачити на прикладі паралельних перекладів “Декамерона” Дж. Бокаччо, де італійський оригінал речення та болгарський відповідник Н. Іванова й Д. Петрова вдвічі довші, ніж український М. Лукаша: *Calandrino, essendogli il vino uscito dal capo, si levò la mattina; e come scese giù guardò e non vide il porco suo e vide l'uscio aperto* || *На следната утрин, когато главата му се избистрила от виното, Каландрино станал, слязъл долу, огледал се и видял, че прасето е изчезнало, а вратата – отворена* || *Прочумався рано-вранці Каландрино після випивки, встав, дивиться – кабана нема, а двері одчинені*). Ми працюємо над

забезпеченням корпусу оригіналами всіх перекладів. Застосування такого підходу дозволяє не тільки побудувати корпус корисного розміру, а й надати базу для дослідження ширшого кола проблем порівняльного мовознавства (через появу опосередкованих відповідників).

Розширення набору мов оригіналів почалося із залучення текстів близькоспорідненими мовами. З'ясувалося, що в мережі найбільше перекладених і українською, і болгарською мовами книг польських та російських, тому поки ми обмежилися цими чотирма слов'янськими мовами. З інших європейських мов було віднайдено і залучено до корпусу паралельні переклади з англійської, німецької, французької, іспанської та італійської мов. Видається цікавим розширити корпус перекладами з інших мов Західної Європи, а можливо, і з неєвропейських, хоча можна передбачити, що із доступністю паралельних текстів буде сутужно.

Очевидно, що чим далі відстань між мовами, тим більшими є розходження між болгарськими й українськими співвідносними перекладами. Але в процесі перекладу з близькоспорідненої мови перекладач під впливом мови оригіналу може свідомо чи несвідомо обрати вираз, близький до наявного в цій мові, але периферійний для мови перекладу. Імовірність цього значно менша для перекладу з неблизькоспорідненої мови, оскільки тотожного виразу в ній, швидше за все, не буде. Виокремити такі лексичні відповідники важче, зате цікавіше, оскільки вони більш “безпосередні”, адже на перекладача менше впливає мова оригіналу, а якщо й впливає, характер цього впливу становить інтерес сам по собі.

Варто зазначити, що тексти, знайдені в мережі Інтернет, часто-густо містять багато помилок розпізнавання, і доводиться їх редагувати зі зверненням до паперового джерела, що потребує немало часу. Наразі триває поповнення корпусу відсканованими книгами паперових бібліотек, хоча низька якість друку часто унеможлиблює розпізнавання тексту, і такий текст виявляється не придатним для корпусних досліджень.

Тексти поділяються на речення за допомогою власних програмних засобів. У випадку деяких текстів з особливо довгими реченнями видається доцільним зарахування крапки з комою, а іноді й двокрапки, до розділових знаків кінця речення. Без цього відповідні одна одній частини тексту можуть виявитися незручно великими, особливо якщо межі речень не збігаються, що трапляється досить часто (до таких текстів можна віднести “Доктор Фаустус” Т. Манна, “Декамерон” Дж. Боккаччо, “Сто років самотності” Г. Маркеса).

Вирівнювання поділених на речення текстів здійснюється за допомогою програми Hunalign [Varga та ін.], <http://mokk.bme.hu/resources/hunalign/>. Часткова автоматизація цього процесу посприяла виявленню проблеми “непаральності” текстів оригіналу та перекладу. Численні розбіжності між зіставлюваними болгарськими й українськими текстами можуть зводитися до скорочення оригінального тексту в перекладі (у трьох основних різновидах: пропуски окремих речень, переформулювання речень у бік скорочення абзацу зі збереженням змісту загалом та викидання з тексту значних його частин: у перекладі П. Кинвою роману П. Загребельного “Переходимо до любові” пропущено вбудовану п’єсу), перестановки цілих абзаців і навіть розділів (так, розділи 2, 3 та 10 роману А. Гуляшки “Пригода опівночі” в перекладі О. Кеткова відповідають розділам 9, 1 та 8) або повної змістової невідповідності частин тексту. Зараз можна тільки здогадуватися, на якому рівні трансформувався текст у кожному випадку: чи на рівні перекладача, який із власної ініціативи намагався “покращити” оригінал, чи то з волі редактора, який мав певні вимоги щодо обсягів перекладу, чи відповідно до загальної настанови щодо спрощення перекладів задля ознайомлення з ними широкої читачької аудиторії. Іншою причиною таких розбіжностей у текстах оригіналу і перекладу може бути наявність різних оригінальних видань, у тому числі й скорочених. Але це вже тема для окремої розвідки.

Такі “недопаральні” тексти викликають неоднозначне ставлення. З одного боку, вони є матеріалом для дослідження, на-

приклад, перекладних трансформацій або історії перекладу, а з іншого – ускладнюють корпусне опрацювання текстів, адже смислові зміни часто супроводжуються формальними перегрупуваннями. Там, де пропускаються або зміщуються речення й абзаци, частини текстів доводиться вирівнювати вручну, а це потребує додаткового часу.

До одиниць, які становлять інтерес для теорії перекладу та не піддаються автоматичному опрацюванню, належать співвідносні, але не паралельні уривки тексту. Наприклад, у романі “Ім’я рози” У. Еко латинський текст Кассінського свідоцтва (X ст.) *Sao ko kelle terre per kelle fini ke ki kontene, trenta anni le possette parte sancti Benedicti* в болгарському перекладі Н. Іванова передано дослівно сучасною мовою як *Знам, че тези земи в границите, в които са очертани, в продължение на трийсет години са били притежание на свети Бенедикт*, а в українському перекладі М. Прокопович його замінено на уривок з “Повчання” Володимира Мономаха (кінець XI – початок XII ст.) *Вскую печална еси, душе моя? Вскую смущаєши мя? Уповаю на Бога, яко исповімся єму* (відповідно до авторського задуму, який полягає саме в байдужості до значення сказаного дійовою особою: для порівняння, в англійському перекладі В. Вівера, 1983, доступному на сайті “Google Книги”, цієї фрази взагалі немає).

Проблеми виникають і з поетичними вставками у прозових текстах: стрункність віршової форми досягається здебільшого за рахунок точності перекладу, тому було вирішено залишати для опрацювання тільки масиви прози. Скороченню може підлягати і цілеспрямовано перекичена мова окремих героїв (прикладом є мова Сальватора зі згаданого вище роману, який в оригіналі і в українському перекладі “говорив усіма мовами і не говорив жодною”, тобто мішанкою слів та конструкцій різних мов; у болгарському перекладі його мова мало відрізняється від літературної).

Хоча й рідко, але трапляються несвідомі розбіжності, спричинені неоднозначністю виразу в оригінальному тексті, який два перекладачі зрозуміли по-різному. Так, польське *Wybiła godzina*

(С. Лем “Фіаско”) передано Д. Андрухівим українською мовою як *Час настав*, а Л. Васілевою болгарською – як *Удари един часът* ‘Пробила перша година’. Спорідненим, частішим (і цікавішим) явищем є неминуча розбіжність у перекладах неоднозначних лексичних одиниць мови оригіналу, як, наприклад, англ. *you* (*ти* чи *ви*) або нім. *Kirsche* (*вишня* чи *черешня*).

Нарешті, у перекладах трапляються неточності та помилки. До неточності можна віднести переклад французького *cochon de lait* ‘молочне порося’ (Ж. Верн “Таємничий острів”) болгарською мовою як *прасенце* (замість *прасе сукалче*): *После с помощта на Наб морякът нагласи ръжсена и добре изкорमेंият кабиай скоро се опеche на пламналия буен огън като обикновено прасенце* (переклад Й. Петрова) || *Потім із Набовою допомогою моряк прилаштував рожен, і майстерно випатрана водосвинка, схожа на молочне порося, незабаром смажилася на веселому ясному вогні* (перекладач – В. Омельченко). Помилку перекладу можна виявити, якщо порівняти паралельні речення *Czas biegu sygnałów nie może być dłuższy od czasu reakcji składników komputera* || *Времето за преминаване на сигнала не трябваше да бъде по-голямо от времето, за което реагират съответните съставни части от компютъра* || *Швидкість руху сигналів не може бути більшою від швидкості реакції складових елементів комп'ютера* (С. Лем “Фіаско”). Чим більша швидкість руху, тим менше часу на нього йде; тож в українському перекладі має бути *швидкість руху сигналів не може бути меншою*.

У процесі накопичення текстів було помічено певну кореляцію між мовами і жанрами творів, яка не сприяє збалансованості корпусу. Навіть якщо знайдено багато текстів, перекладених з якоїсь мови, це може означати, що є багато перекладів творів одного автора. А таке не бажано з погляду статистичних досліджень лексики. В окремих випадках у найчастотніші слова потрапляють вигадані авторами власні назви (наприклад, із фантастичних романів І. Єфремова), що зовсім зайве. Коли авторів двоє-троє, з великою імовірністю можна припускати, що вони працювали в одному жанрі. Тож зараз у болгарському секторі

переважає політичний детектив (А. Гуляшки, Б. Райнов), у польському – історичний роман (Б. Прус, Г. Сенкевич), у російському – наукова фантастика (О. Беляєв, І. Єфремов). Дається взнаки і мала кількість перекладачів (на стиль перекладу впливають їхні індивідуальні вподобання, особливо в лексиці та фразеології). Зрозуміло, що буде нелегко вийти за межі белетристики і домогтися того, щоб у корпусі були наявні паралельні українсько-болгарські публіцистичні, наукові, мистецтвознавчі, мемуарно-біографічні та інші тексти в представницькій кількості, але слід прагнути до зменшення згаданого дисбалансу хоча б у межах художніх жанрів.

#### **Корпусне дослідження лексичних перекладних еквівалентів**

Паралельний корпус надає широкі можливості для статистичного вивчення міжмовних лексичних відповідностей з метою уточнення значень слів або кореляцій між словами у певних значеннях та умовами використання цих слів. Послідовність кроків роботи з лексичними еквівалентами в паралельному корпусі може бути такою:

- окреслюємо в одній з двох мов лексико-семантичне поле (ЛСП), лексико-семантичну групу (ЛСГ) чи окремі лексико-семантичні варіанти (ЛСВ), особливості перекладу яких нас цікавлять. Відбір одиниць ЛСП або ЛСГ добре здійснити за допомогою ідеографічного словника, якщо це можливо для досліджуваної мови. Також доцільно уточнити лексичні значення відібраних слів за тлумачним словником;
- робимо пошук ЛСВ у корпусі й отримуємо певну кількість пар речень, що містять якесь із шуканих слів у секторі, відповідному вихідній мові, разом з інформацією про те, у якому тексті вони знайшлися;
- відфільтруємо омоніми, залишаючи словоформи з потрібними нам лексичними значеннями;
- визначаємо перекладні відповідники словоформ у паралельних реченнях;
- класифікуємо перекладні відповідники й аналізуємо їх за якісними та кількісними параметрами.

F1241	▶ ◀ ✕ ✓ ⚡	Но із бях упорит и получих апартаментче два етажа по-долу. ~~~~ С бая и с всичките други удобства, но
E	F	G
1235	Vg_AG_AZ-1	Самолетът се падна във висините, проби камаршата мъгли и см. Митак заер нося вгору, пробив громадама злар, і вже за кілька секунд
1236	Vg_AG_AZ-1	Старши милиционерът е ударен с пистолет по главата само тридесет Старши милиционерът взривоно пистолетом по голови через тридцять секунд
1237	Vg_AG_AZ-1	Очите му бяха свързани хищноизгряни от фосфоресциращи стрел. Очи. Авакумов: бяхи принити до святих стрілок, а сердце, здається, секунд
1238	Vg_AG_AZ-1	В следната секунда той оръяна въз вътрешния джоб на шифера с За хвилин Авакум дръств із внутрішньої кишені плаща и ятиям секунд
1239	Vg_AG_AZ-1	По този начин той си осигурява няколко часа спокойствие и свобода. Тяма чинюм, він гарантує собі кілька годин безпекі і свободи. часа
1240	Vg_AG_AZ-1	Съобразяването на тия пет часа, когато си прекарал със Сид, е пог. — А прозав ти дї ять Годин із Сією приблизно так, — прозав. часа,
1241	Vg_AG_AZ-1	Но аз бях упорит и получих апартаментче два етажа по-долу. ~~~~ А
1242	Vg_AG_AZ-1	Първо — и най-неопитното око, шом погледне тия островърхли гот. По-перше, і найнеосвідченіша людина, гланувши на кімчати. часа
1243	Vg_AG_AZ-1	Сините в мъглята му подказваха още две интересни неща. Липшего Х Сині ворсинки підказали йому ще дві цікаві обставини: невдало. часа
1244	Vg_AG_AZ-1	Май че на два пъти е плавала, но за по няколко часа. — Нобито приєднала дачи, та ненадовго, на кілька годин. часа.
1245	Vg_AG_AZ-1	Погледна часовника си — наблюдаваше два часаът сама полунощи. Подлившись на годинник — наблюдавался арта. голина ноци. часа.
1246	Vg_AG_AZ-1	Наблюдаваше четри часаът, но мразкът беше все така дълбок, като наблюдаваше четвърта година рану, але темрява була густа, и. часа.
1247	Vg_AG_AZ-1	Беше десет часаът преди обяд. Була десята година рану. часа.
1248	Vg_AG_AZ-1	Наблюдаваше осем часаът. Наблюдавался восьма година. часа.
1249	Vg_AG_AZ-1	Пет часаът — първоначално шестима, а после четворка. Пята година. ~~~~ Спочату лїстий номер травлаи, потім че. часа.
1250	Vg_AG_AZ-1	Или ли десет часа?
1251	Vg_AG_AZ-1	— Колако е часаът?
1252	Vg_AG_AZ-1	Ще ти кажат, че съм излезла между седем и осем часа. Вони скажуть тобі, що я вийшла між сьомою та восьмою год. часа.
1253	Vg_AG_AZ-1	А свидетелът Марко Крумов, който живее в източните покрайнини, А свидет. Марко Крумов, який живє на східній околиці села, ств. часа

Рис. 1. Вигляд вікна корпусу в процесі дослідження слів часу



Усю послідовність дій можна повторити для іншої мови. Цікавим моментом є аналіз кореляції між частотою пар перекладних еквівалентів та мовою оригіналу, оскільки “навіть коли лексичну одиницю цільової мови можна вжити як перекладний відповідник леми вихідної мови (що не завжди можливо), перекладом цієї лексичної одиниці не завжди є лема вихідної мови оригіналу” [Лендау : 314].

В іншій постановці дослідження ЛСП, ЛСГ чи ЛСВ визначаються одразу в обох мовах, і в корпусі шукаються пари паралельних речень, що містять відібрані одиниці в болгарській, українській чи обох частинах корпусу.

Корпусний аналіз перекладних відповідників можна робити для будь-яких лексико-семантичних одиниць. Ми здійснили його на робочій версії корпусу текстів зі слов’янськими оригіналами (українська частина налічує приблизно 2½ млн. слововживань) для групи болгарських і українських іменників на позначення часу [Держански; Derzhanski, Siruk]. На рис. 1 показано вигляд робочого вікна дослідження, де у стовпчиках вказується порядковий номер пари речень, код джерела, болгарське та українське паралельні речення. Вгорі можна прочитати повний текст виділеного під номером 1241 у стовпчику F болгарського речення. Верхнім штрихом перед початком позначаються слова, які шукає дослідник.

Дослідження лексем на позначення часу дозволило зробити деякі висновки щодо тривалості визначених цими лексемами періодів. Так, лексема *момент* позначає точку (або дуже короткий сегмент) на часовій осі; чим довшим стає сегмент, тим більше імовірності, що він називатиметься не *момент*, а буде визначатися як болгарське *хвилина* чи українське *хвилина*. Щодо *миг* та *мить*, то ці слова займають середню позицію. Це також пояснює відносно рідке вживання “екстремальних” пар перекладних еквівалентів *хвилина* : *момент* і *момент* : *хвилина* у порівнянні з парами *хвилина* : *мить*, *миг* : *момент*, *момент* : *мить* і *миг* : *хвилина*. Слово *хвилина* в українській мові вживається на позначення невизначеного короткого проміжку часу активніше, ніж його ос-

новний болгарський відповідник *минута*. Це може мати етимологічне пояснення: *минута* – порівняно нове слово, яке від початку позначає часовий відтинок з точно визначеною і цілком усвідомлюваною тривалістю, тоді як *хвилина* < “*хвиля* ‘хвилина; короткий відрізок часу, мить’ є давнім запозиченням з німецької через посередництво польської; нім. *Weile* ‘певний час’ < двн. *hwīl, hwīl(a)* ‘час, година’ [...] припускається спорідненість *хви́ля* (час) і *хви́ля* (на воді)” [Мельничук : 166]. Слово *момент* вживається частіше (тобто глибше засвоєне) в болгарській мові, ніж в українській. Попередні експерименти над близьким за розміром корпусом текстів з англійськими, німецькими, французькими, італійськими й іспанськими оригіналами дають подібні результати.

Цікавим є також дослідження перекладних еквівалентів, які не становлять класичної пари “іменник – іменник”. У корпусі іменники на позначення часу та конструкції з ними можуть перекладатися:

- іменником, який не має значення часу: *В подобни мигове не е полезно да се проявява нетърпение || В таких випадках не треба гарячкувати* (Б. Райнов “Пан Ніхто”);

- прислівником часу: *и в такива мигове пейзажът добива нещо от двусмисления израз на човек || й тоді краєвид набував чогось від двозначної гримаси людини* (Б. Райнов “Велика нудьга”); при цьому семантика темпоральності може підкреслюватися вживанням відповідного дієслова зі значенням роду дії: *Най-напред блесна малко, ослепително слънце, което в следните мигове сякаш изпълни всички илюминатори || Спочатку спалахнуло невелике сліпуче сонце і відразу засяяло в усіх ілюмінаторах* (П. Вежінов “Коли ти в човні...”);

- прислівником, який не має значення часу: *Имал съм и по-тежки мигове || Бувало й важче* (Б. Райнов “Велика нудьга”). Форма вищого ступеня порівняння українського прислівника *важче* мотивована аналогічною формою болгарського прикметника *по-тежки*. Семантика множинності не втрачається, а переходить з іменника у формі множини *мигове* до дієслова із семантикою багатократності *бувало*.

### **Корпусне знаходження лексичних перекладних відповідників**

Ще одне цікаве застосування паралельного корпусу полягає в автоматичному віднайденні лексичних перекладних еквівалентів, кульмінацією якого є автоматизована побудова двомовного словника. Ця процедура ґрунтується на пошуку пар слів, які найчастіше трапляються у співвідносних реченнях паралельного корпусу.

Якість такого словника краща, коли корпус великий, хоча це збільшує вимоги до обчислювальних ресурсів. До факторів, через які якість словника погіршується, належать:

- неточність перекладу, характерна для художнього тексту (особливо для секторів, які містять переклади з третіх мов);
- велика довжина речень, властива деяким авторам, а також створювана автоматично через розходження меж речень у пари вирівнюваних текстів;
- істотна несхожість граматичної будови мов.

З огляду на розвинену словозміну та розгалуженість граматичних парадигм як української, так і болгарської мов кращим буде застосування граматичного словника (лематизатора) на початковому етапі автоматичного визначення лексичних відповідників, для того щоб пошук здійснювався не в просторі пар словоформ, а в істотно вужчому просторі пар лексем.

### **Розвиток проекту**

Роботу над корпусом болгарських та українських паралельних текстів передбачається продовжувати за трьома основними напрямками.

По-перше, розвиватимуться кількісний та якісний склад корпусу. Мова йде як про додавання нових секторів (мов-оригіналів), поповнення корпусу новими творами та їхніми перекладами (у широкому розумінні паралельності текстів), так і про жанрове розширення корпусу, тобто збільшення різноманітності жанрів та зменшення кореляції між мовами і жанрами.

По-друге, доопрацювання вимагає якості текстового матеріалу та його представлення. Це стосується широкого діапазону завдань від пошуку і виправлення технічних помилок до забез-

печення текстів морфологічною та синтаксичною розмітками, виконаних відповідно до прийнятих стандартів.

По-третє, потрібно розробити супровідні допоміжні програмні засоби, а саме лематизатор (на основі граматичних словників) та пошукову систему, і створити онлайн-версію корпусу.

1. *Гарабик Р.* Параллельный русско-словацкий корпус / Р. Гарабик, В. П. Захаров // Труды Международной конференции “Корпусная лингвистика – 2006”. – СПб., 2006. – С. 81–87.
2. *Демська О.* Текстовий корпус: ідея іншої форми / О. Демська. – К., 2011.
3. *Держански И.* Думите за време в български и украински език (върху материал от паралелни текстове) (у друці) / И. Держански.
4. *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., 2005. – С. 263–296.
5. Етимологічний словник української мови : у 7 т. / редкол.: О. С. Мельничук (голов.ред.) [та ін.]. – Т. 6 : У – Я / уклад. Г. П. Півторак [та ін.]. – К., 2012.
6. *Лендау С. І.* Словники: мистецтво та ремесло лексикографії / С. І. Лендау. – К., 2012.
7. *Parallel Texts. Using Translational Equivalents in Linguistic Typology.* Theme issue in Sprachtypologie and Universalienforschung STUF 60.2. / eds. *M. Cysouw, B. Wälchli.* – 2007.
8. *Derzhanski I.* Brief Time Words in Bulgarian and Ukrainian (Using Evidence from Parallel Texts) / *I. Derzhanski, O. Siruk* // The Eight International Conference “Formal Approaches to South Slavic and Balkan Languages”: Book of Abstracts. – Zagreb, 2012. – С. 12.
9. *Varga D.* Parallel Corpora for Medium Density Languages / D. Varga, L. Németh, P. Halácsy, A. Kornai, V. Trón, V. Nagy // Proceedings of the RANLP. – 2005. – P. 590–596.
10. *Vitas D.* Preparation and Exploitation of Bilingual Texts / D. Vitas, C. Krstev, E. Laporte // Lux Coreana. – № 1. – Han-Seine, 2006. – P. 110–132.

*Речь идет о создании корпуса болгарских и украинских параллельных текстов и его использовании в исследовании лексической семантики украинских и болгарских переводных эквивалентов.*

**Ключевые слова:** корпусная лингвистика, параллельные тексты, корпус, лексическая семантика, переводные эквиваленты.

*The paper introduces the Corpus of Bulgarian and Ukrainian Parallel Texts and discusses its use for the analysis of the lexical semantics of Bulgarian and Ukrainian translation equivalents.*

**Keywords:** corpus linguistics, parallel texts, corpus, lexical semantics, translation equivalents.

**Стаття надійшла до редакції 4.09.2012**