

3. Сербенська О. А. Актуальне інтерв'ю з мовознавцем: 140 запитань і відповідей / Сербенська О. А., Волощак М. Й. – К. : Просвіта, 2001. – С. 108.
4. Українська мова : Енциклопедія. – К. : Просвіта, 2000. – С. 377.

Темчишин Є.А.

магістрант,

Національний університет «Львівська політехніка»

КВАНТИТАТИВНИЙ АНАЛІЗ ПАРАЛЕЛЬНОГО КОРПУСУ ПОВІСТІ ЕРНЕСТА ГЕМІНГВЕЯ «СТАРИЙ І МОРЕ»

Для лінгвістичних досліджень другої половини ХХ ст. – початку ХХІ ст. властиві поява та стрімкий розвиток корпусної лінгвістики, яка надає такий інструмент дослідження, як лінгвістичні корпуси текстів, що слугують потужною інформаційно-лінгвістичною підтримкою для вивчення різних аспектів мови. У сучасній лінгвістичній науці існують різні визначення текстового корпусу.

Йдучи за працями О. Демської [3], І. Кульчицького [5; 6], У. Н. Фрэнсиса [9], В. Широкова [4], під текстовим корпусом розуміємо призначену для різноманітних досліджень лінгвістично компетентно описану сукупність мовних чи мовленнєвих даних, поданих в електронній формі та споряджених відповідним спеціалізованим програмним забезпеченням. Типологія корпусів текстів є узвичаєною, представлена в усіх роботах з корпусної лінгвістики. Для теорії та практики перекладу особливе значення мають автоматизовані корпуси паралельних текстів, тобто оригінальних текстів та їх перекладів.

Паралельні корпуси текстів – це потужний ресурс для міжмовних досліджень у галузях контрастивної лінгвістики, перекладознавства, лінгвістичної типології, стилістики, лінгвістики тексту, тощо. Важливу роль вони відіграють у вивченні ідіолекту письменника і способів передачі його рис у перекладі.

У нашому дослідженні ми приділяємо першочергову увагу лексичному складу оригіналу та перекладу, а отже за допомогою автоматичного опрацювання паралельного корпусу оригіналу повісті Ернеста Гемінгвея «Старий і море» [10] та її перекладу українською

мовою Володимиром Митрофановим [2], а також статистичних обрахунків, встановимо ряд важливих характеристик, які можуть стати основою для з'ясування рис ідіостилу письменника, допоможуть зробити висновки про естетичну значимість текстів оригіналу та перекладу та їх художню відповідність.

З погляду квантитативно-лінгвістичного аналізу текстів можна виділити ряд актуальних проблем, пов'язаних із стилістичним аспектом дослідження лексики даного тексту [1], зокрема питання про обсяг тексту – загальну кількість слів у тексті (N), обсяг словника слівформ, кількість слівформ у тексті (V_f), обсяг словника лексем, кількість лематизованих слів у тексті (V), Нарах *legomena* (V_1) – слова, які трапилися у досліджуваній вибірці один раз, тобто мають частоту 1, кількість самостійних частин мови (N_{ps}), кількість службових частин мови (N_{psa}). На основі цих статистичних даних можна обчислити:

- багатство словника, індекс різноманітності (I_d) – відношення обсягу словника лексем до обсягу тексту;
- середню повторюваність слова у тексті (I_{wr}) – відношення обсягу тексту до обсягу словника лексем;
- індекс винятковості – обчислюється окремо для словника і для тексту, він характеризує варіативність лексики, тобто частку тексту (словника), яку займають слова, що трапилися 1 раз: індекс винятковості для словника (I_{ev}) – відношення кількості лексем із частотою 1 до загальної кількості лексем; індекс винятковості для тексту (I_{en}) – відношення кількості лексем із частотою 1 (V_1) до обсягу тексту (N);
- індекс концентрації словника (I_{vc}) – відношення кількості слів у словнику з абсолютною частотою 10 і більше (V_{10}) до загальної кількості слів у словнику;
- індекс лексичної щільності (I_d) – відношення самостійних частин мови (N_{ps}) в тексті до загальної кількості слів;
- іменних означень (I_{nat}), тобто індекс епітетизації (відношення суми вживань іменників (V_n) до суми вживань прикметників (V_{adj});
- індекс дієслівних означень (I_{vat}) – відношення суми вживань прислівників (V_{adv}) до суми вживань дієслів (V_v);
- ступінь номінальності (I_{nom}) – відношення суми вживань іменників (V_n) до суми вживань дієслів (V_v).

Аналіз повісті Ернеста Гемінгвея «Старий і море», проведений на матеріалі електронного маркованого паралельного корпусу, охоплює ряд характеристик, співставлення котрих наведено у таблиці 1.

Таблиця 1

Статистичні показники оригіналу та перекладу

Статистичний показник	Оригінал	Переклад
Обсяг тексту	26523	22548
Кількість різних словоформ	2740	6007
Кількість різних слів	1915	3231
Нарах leqomena:текст	1314	3911
Нарах leqomena:словник	825	1582
Кількість слів з частотою ≥ 10	344	335
Кількість самостійних частин мови: текст	2638	5823
Кількість самостійних частин мови: словник	1853	3108
Кількість службових частин мови: текст	91	184
Кількість службових частин мови: словник	62	123
Середня повторюваність слова у тексті:	13,85	6,98

На основі наведених статистичних даних були обчислені індекси, що характеризують лексичний рівень оригіналу та перекладу повісті. Результати обчислень проілюстровано діаграмою гістограмою (рис. 1).

При вивченні показників лінгвістичних характеристик велике значення має наступна постановка питання: деякий лінгвістичний показник визначений по двох різних вибірках. Як правило, значення цього показника в одній вибірці відрізняється від його значення в іншій вибірці. Чи істотно це розходження?

У тих випадках, коли закон розподілу цього показника близький до нормального, відповідь на поставлене питання дає критерій Стюдента. Вимогу нормальності можна опустити, якщо об'єми порівнюваних вибірок досить великі (більше 30). Критерій Стюдента t можна обчислити, скориставшись в Excel функцією СТЬЮДЕНТ.ТЕСТ.

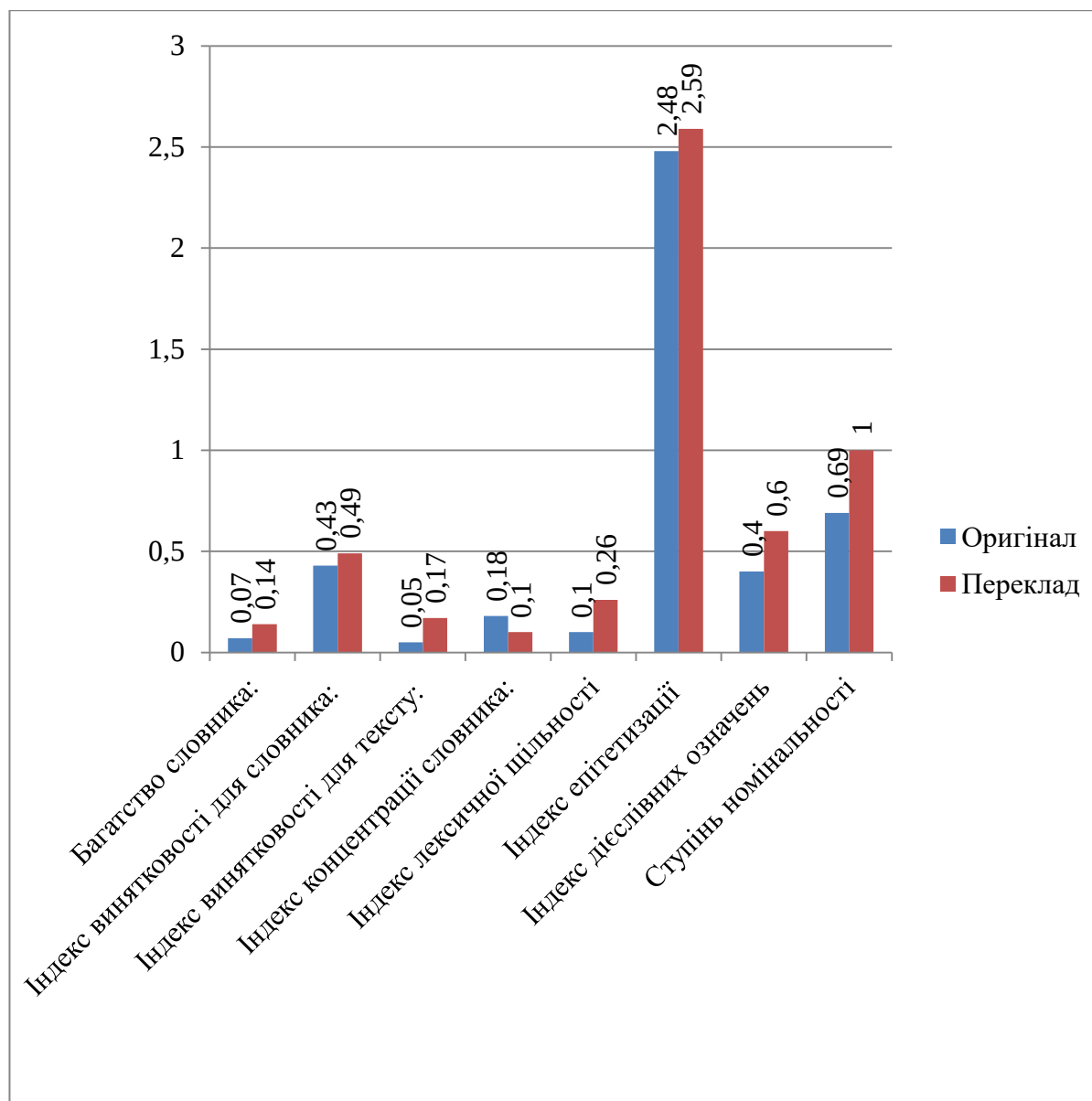


Рис. 1. Квантитативні характеристики лексичного рівню оригіналу та перекладу

Джерело: розроблено автором

Щоб вирішити, чи свідчить *t*-критерій про істотне розходження, треба оцінити його за таблицею критичних значень *t* [8]. Оцінку цю проводять за допомогою визначення кількості ступенів свободи, яке у нашому випадку $f=8-2=6$. Розходження вважається істотним, якщо обчислене значення *t* більше від табличного для заданого рівня значущості. У нашому випадку 0,92 менше від найменшого числа у ряду. Це значить, що розходження у статистичних показниках оригіналу і перекладу неістотне і статистично допустиме.

Зауважимо, що статистичні показники не дають підстав для визначення естетичних можливостей тексту перекладу. Статистичні

характеристики відображають лише деякі структурні особливості текстів і можуть лише вказувати на приховані від простого спостереження особливості індивідуального чи функціонального стилю [7, с. 152]. Якісний аналіз одержаних даних можна зробити у межах лексико-семантичного дослідження домінант ідіостилу твору та їх відповідників у перекладі.

Список використаних джерел:

1. Бук С. Інтерпретація лексики роману Івана Франка «Великий шум» у кількісно-статистичному аспекті / с. Бук // *Studia Metodologica*. – Вип. 30. – Тернопіль : ТНПУ ім. В. Гнатюка, 2010. – С. 145-153.
2. Гемінгвей Е. Старий і море / Е. Гемінгвей / Львів: Видавництво Старого Лева, 2017. – 104 с.
3. Демська О. Текстовий корпус: ідея іншої форми / О. Демська. – К.: ВПЦ НаУКМА, 2011. – 282 с.
4. Корпусна лінгвістика / За ред. В. А. Широкова. – К.: Довіра, 2005. – 471 с.
5. Кульчицький І. М. Деякі аспекти створення та використання паралельних корпусів / Данчевська Ю. О., Кульчицький І. М., Ліхнякевич І. О. // *Науковий вісник ВНУ ім. Лесі Українки. Серія: Філологічні науки*. – Луцьк, 2013. – С. 48–52.
6. Кульчицький І. Технологічні аспекти укладання корпусів текстів / І. М. Кульчицький // *Дані текстових корпусів у лінгвістичних дослідженнях: Монографія* / За ред. О. П. Левченко. – Львів: Вид-во Львівської політехніки, 2015. – С. 29–45.
7. Перебийніс В. С. Частотні словники та їх використання / В. С. Перебийніс, М. П. Муравицька, Н. П. Дарчук. – К.: Наукова думка, 1985. – 204 с.
8. Таблица распределения Стиюдента // <http://old.exponenta.ru/educat/referat/XIkonkurs/student5/tabt-st.pdf>
9. Фрэнсис У. Н. Проблемы формирования и машинного представления большого корпуса текстов / Пер. с англ. И. Г. Сабуровой // *Новое в зарубежной лингвистике*. – Вып. XIV. – М.; Прогресс, 1983. – С. 334–352.
10. Hemingway E. The old man and the sea [Electronic resource]. / E. Hemingway – NY.: Charles Scribner's Sons, 1952. – 127 p. Access mode: https://la.utexas.edu/users/jmciver/Honors/Fiction%25202013/Hemmingway_The%2520Old%2520Man%2520and%2520the%2520Sea_1952.pdf