**Міністерство освіти й науки України**

**Львівський національний університет імені Івана Франка**

Факультет електроніки та комп’ютерних технологій

*з предмета: Комп‘ютерна лінгвістика*

Звіт

про виконання лабораторної роботи № 10

**«Дослідження базових лінгвістичних властивостей рандомних символьних**

**послідовностей»**

Виконав:

Студент групи

Фес-32с

Бойко Кирило

Львів 2024

***Завдання***

1. Для досліджень використати рандомні та рандомізовані тексти, генеровані в

лабораторній роботі №9.

2. Дослідити усереднений по всіх словах параметр R кластеризації слів у цих

текстах і усереднений по всіх словах параметр γ кореляцій слів у них.

3. Дослідити ті ж параметри для такої ж кількості природних текстів, написаних

різними мовами.

4. Зробити висновки стосовно схожості або відмінності параметрів R і γ для

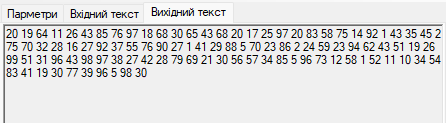
природних і рандомних або рандомізованих текстів.

**Хід виконання роботи**

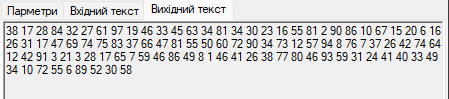
Запустив програму **+shuffler (words only)** та завантажив до неї перший текст

***Clemencia Novela de costumbres by Fernán Caballero***

Провів перевірку, кількість циклів рандомізації N = 50 000



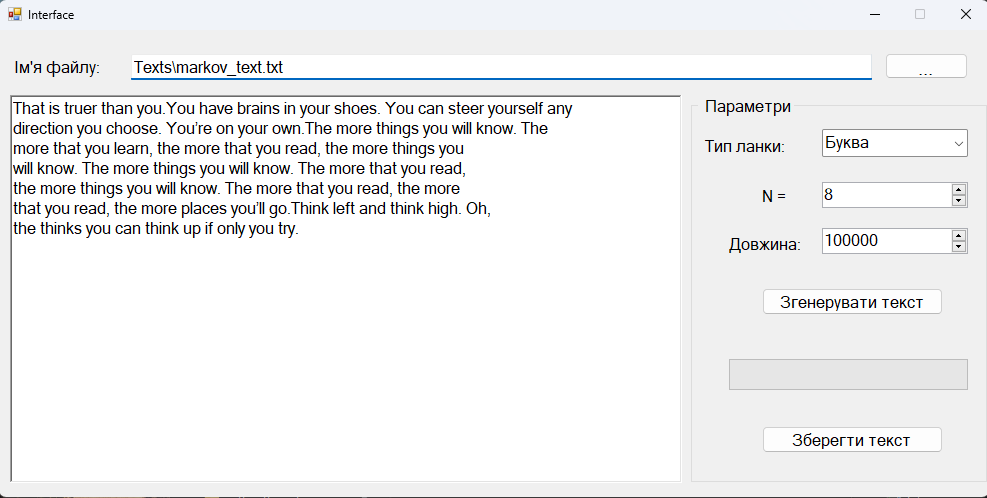
Завантажив текст мавпи Міллера з попередньої лабораторної роботи



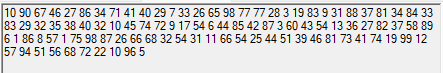
Відкрив програму **+Markov texts** з лабораторної роботи 27 та згенерував текст, з такими налаштуваннями:

*N = 8  
Тип ланки: букви*

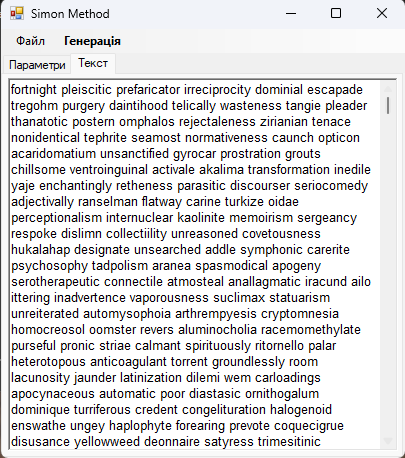
*Довжина: 100000*

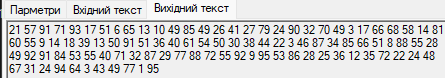


Зберіг цей текст і завантажив його в поточну програму лабораторної роботи для подальшого аналізу

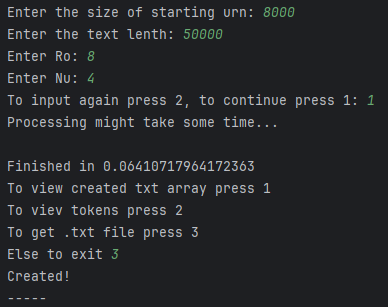


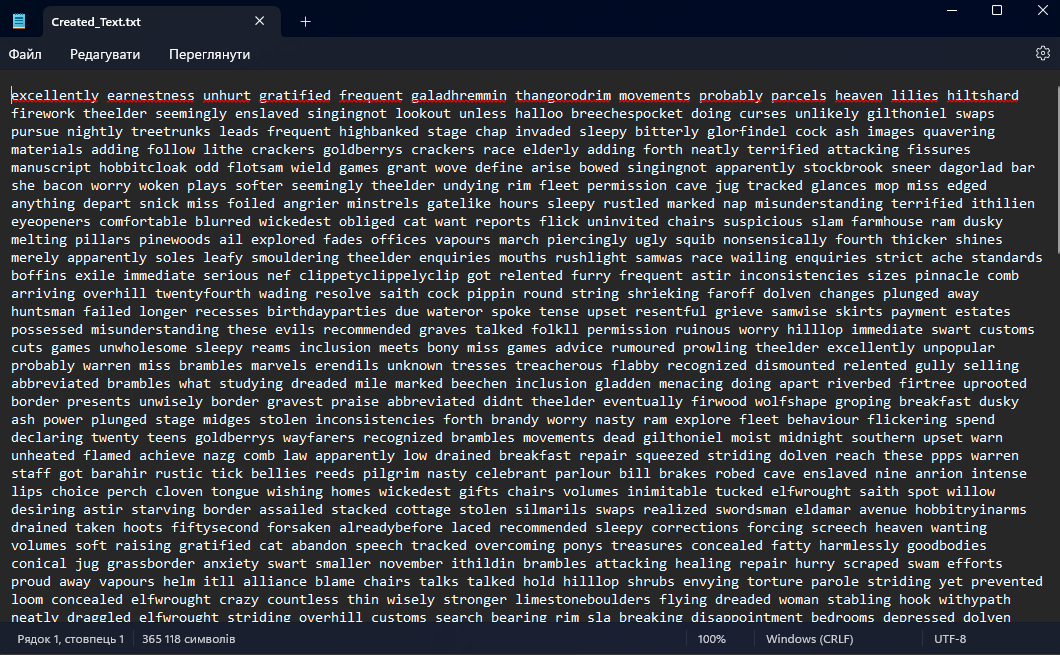
Згенерував новий текст в програмі **+Simon\_3in1**

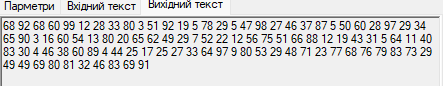


**Результат дослідження:**  
  


Згенерував текст за допомогою програми **+PolyaTexts** і завантажив його для подальшого дослідження

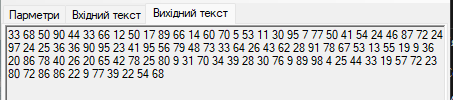


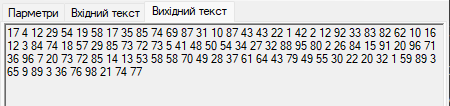
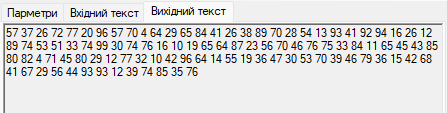


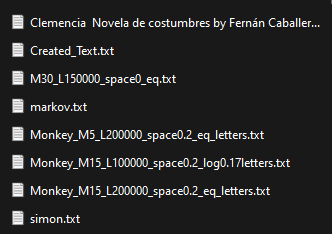


Далі, я згенерувала ще 3 текста за допомогою програми **+monkey\_texts**

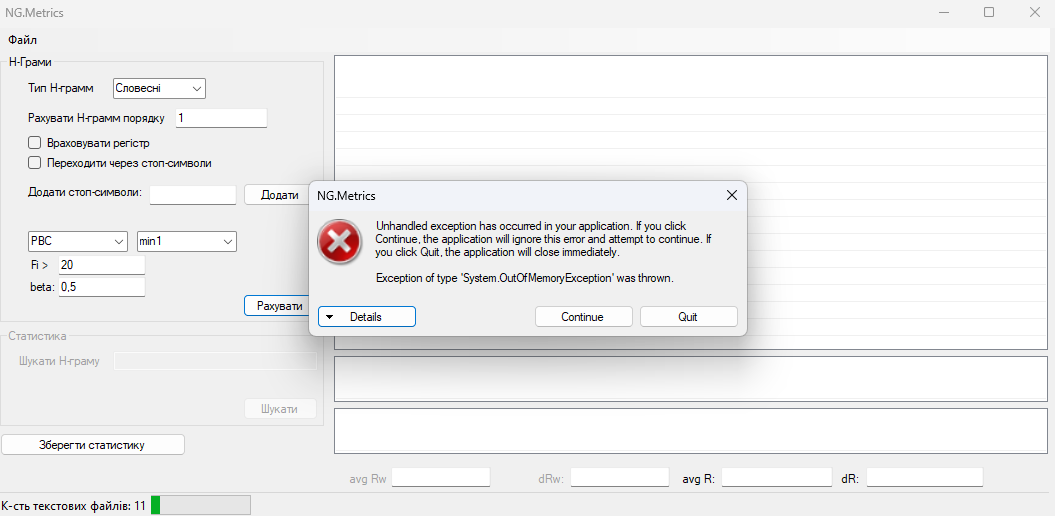
Та провела дослідження з ними:







На жаль, мені не вдалося виконати наступні завдання, оскільки програма давала збій у роботі



**Висновок**: У процесі виконання роботи було досліджено лінгвістичні характеристики рандомних та рандомізованих текстів за допомогою параметрів кластеризації RRR та кореляції γ. Для аналізу були згенеровані тексти за допомогою різних моделей, таких як текст мавпи Міллера, тексти Маркова, а також моделі Саймона і Полі. Отримані результати показали, що параметри кластеризації та кореляції в природних текстах суттєво відрізняються від тих, що характерні для рандомних текстів. Це підтверджує можливість використання цих параметрів для розрізнення різних типів текстів, що є важливим інструментом для лінгвістичних досліджень та аналізу текстової інформації.