

DOI: <https://doi.org/10.32839/2304-5809/2019-11-75-209>

УДК 811.161.2'324'373'38:8-31

Огородник В.В., Лотоцька Н.Я., Кульчицький І.М.
Національний університет «Львівська політехніка»

ЛІНГВОСТАТИСТИЧНИЙ АНАЛІЗ ПОВІСТЕЙ РОМАНА ІВАНИЧУКА

Анотація. В даній статті досліджено використання квантитативного методу в галузі лінгвістики. Розраховано довжину слів та розділів у повістях Романа Іваничука, а саме «Соло на флейті», «Неспокутне», «Лісова повість», «Погоня за Пегасом», «І земля, і зело, і пісня». Більшість досліджень творчості Романа Іваничук полягали у висвітленні жанрового аспекту творів автора. Кожен видатний автор, який є представником лінгвокультурного ареалу, демонструє унікальний стиль у своїх творах та виробляє свій неповторний почерк. Індивідуальний стиль автора кожен з вчених тлумачить по-своєму, але на основі численних досліджень можна стверджувати, що це унікальний «почерк» кожного письменника, своєрідний вияв творчості, прояв манери написання та прагнення втілити певні думки. На формування індивідуального стилю автора впливають також політичні рухи, тенденції розвитку літератури, певний колорит.

Ключові слова: ідіолект, ідіостиль, статистичний метод, квантитативна лінгвістика, довжина слова.

Ohorodnyk Viktoriia, Lototska Nataliia, Kulchytskyi Ihor
Lviv Polytechnic National University

LINGUISTIC STATISTICAL ANALYSIS OF NOVEL BY IVAN IVANYCHUK

Summary. This article explores the use of the quantitative method in the field of linguistics. The length of the words and chapters in the novels by Roman Ivanychuk is calculated. Linguistic science is rich in research methods and approaches to text, but recently there has been a tendency to use quantitative methods for the needs of linguistics. It is believed that accurate results can be obtained at the junction of these two sciences. Many researches in the early twentieth century turned to the use of statistical methods for linguistic research. Statistical methods that significantly modify and increase our knowledge of both the language system itself and the features of its functioning are used in linguistics in three conventional directions: various quantitative calculations, construction of stochastic linguistic models using the data obtained in the previous direction, verification by statistical methods of various hypotheses about one or another linguistic phenomenon. Therefore, quantitative studies of the Ukrainian language remain relevant today. Most of the studies of Roman Ivanychuk's work consisted in covering the genre aspect of the author's works. Each prominent author, who is representative of the linguistic and cultural area, demonstrates a unique style in his works and produces his own unique handwriting. Each scientist interprets the individual style of the author in his own way, but on the basis of numerous researches we can claim that this is a unique "handwriting" of each writer, a peculiar expression of creativity, a manifestation of the manner of writing and the desire to embody certain thoughts. The formation of the individual style of the author is also influenced by political movements, trends in the development of literature, a certain flavor. The author's uniqueness can be identified using the structure of works, the main subject. The developed style demonstrates the artist's attempts to find himself, the search for something new, the desire to convey certain ideas to readers. Using statistical analysis, we can trace the features of the author's creative activity, investigate the frequency of grammatical and lexical units.

Keywords: idiolect, idiostyle, statistical method, quantitative linguistics, word length.

Постановка проблеми. Тенденція використання квантитативних методів досліджень у лінгвістичній науці щоразу зростає. Статистику використовують не лише в математиці, але й лінгвістиці, зокрема для вивчення особливостей авторської мови [7, с. 139]. Крім цього, у мовознавстві використовують такі статистичні підходи, як «хі-квадратичний, варіативний, кластерний, факторний аналіз та аналіз головних компонентів»[4].

Застосування кількісного критерію в лінгвістичних дослідженнях відоме здавна. Такі мовознавчі поняття, як фонетичний закон, продуктивність морфем, критерій спорідненості мов тощо базувалися в основному на квантитативних характеристиках. Звукові еквіваленти, які нерідко фіксують в мові чи декількох споріднених мовах, є закономірними. Морфема, котрі часто використовують для творення нових словоформ, є продуктивними. Незважаючи на те, що раніше математичний критерій використовували стихійно, зараз його застосовують свідомо і цілеспрямовано.

Статистичні методи широко використовують у мовознавстві. Наприклад, у лексикографії ста-

тистику використовують для складання частотних словників, які допомагають визначити найвживанішу лексику. До того ж статистичні методи пов'язані зі розробленням методики навчання іноземної мови та оптимізації навчання [8, с. 211].

Статистичні методи відіграють велику роль у прикладній лінгвістиці (для автоматичного опрацювання природної мови та мовлення, для створення систем машинного перекладу та електронних словників), психолінгвістиці, соціолінгвістиці, порівняльно-історичному мовознавстві, прагмалінгвістиці тощо. За допомогою статистики мовознавство вирішує широке коло завдань. Використання статистичних методів уможливило встановлення семантичної близькості слів, визначає структуру семантичного поля, дослідження стилю письменника та атрибуції текстів, виявлення хронології мовних явищ [3, с. 66].

Однією із закономірностей мовних одиниць у функціональних стилях, з погляду статистики, є засада строгої кількісної залежності між абстрактністю текстового змісту, що переважає в ньому, типом мислення автора та абстрактністю у мовленні, тобто між

частотним використанням слів і граматичних форм з їх лексико-граматичним значенням [7, с. 141].

Отже, для досліджень у сфері мовознавства використовують імовірно-статистичні методи математичної лінгвістики, а також методи та прийоми математичної теорії множин. До них відносять моделювання множин, визначення множин, взаємне накладання множин для їхнього порівняння.

Порівняльні методи широко застосовують для вивчення тематичних та формальних класів мовних одиниць і, хоч кількісні методи, зазвичай, зараховують до практичної галузі, їх досить часто застосовують для вирішення теоретичних лінгвістичних проблем.

Аналіз останніх досліджень і публікацій.

Серед українських дослідників в галузі лінгвостатистики є С. Бук [2], яка розкрила та описала розвиток та становлення статистичної лінгвістики як науки, виклала основні поняття та категорійний апарат статистичної лінгвістики, представила основні кількісні характеристики одиниць різних рівнів мови та мовлення, статистичну структуру тексту, розглянула найважливіші проблеми стилеметрії, глоттохронології, авторської та часової атрибуції тексту. Також теоретичні та прикладні проблеми структурно-математичної лінгвістики вивчав В.С. Перебийніс [9]. Проблеми квантитативної лінгвістики описано в праці Альтман Г. «Мода та істина в лінгвістиці» [1].

Методи математичної та структурної лінгвістики надають точні дані для вивчення лінгвістичних об'єктів і явищ та відповідають на запитання, які лише на рівні мовознавчої компетенції неможливо з'ясувати. Лінгвістичні моделі будують на основі отриманих числових даних, які наочно розкривають суть досліджуваних явищ. Статистичні методи застосовують для моделювання статистичних структур поетичних текстів та визначення ступеня дії фактору авторської манери письма. Так, Г. Альтман, К.Б. Бектаев, Г.Я. Мартиненко, Р.Г. Піотровський, С.В. Чебанов провели низку досліджень, в яких здійснено підрахунок і вимірювання стилістичних явищ з метою упорядкування текстів та їх частин [10, с. 350].

Величною постаттю української літератури є письменник, громадський діяч, один із засновників Наукового товариства імені Тараса Шевченка – Роман Іваничук. Читаючи твори митця, можна відкрити для себе нові звичаї та традиції українського народу, особливості етносу, побачити мовний досвід письменника, відчути багатство української мови. Романа Іваничука можна вважати справжнім май-

стром історичного роману, адже він показує своїм читачам незвідані сторінки нашої історії.

Кожен видатний автор, який є представником лінгвокультурного ареалу, демонструє унікальний стиль у своїх творах та виробляє свій неповторний почерк. На формування індивідуального стилю автора впливають також політичні рухи, тенденції розвитку літератури, певний колорит. Авторську унікальність можна виявити за допомогою структури творів, основної тематики. Вироблений стиль демонструє спроби митця знайти самого себе, пошуки чогось нового, бажання донести до читачів певні ідеї. За допомогою статистичного аналізу можна прослідкувати особливості творчої діяльності автора, дослідити частотність граматичних і лексичних одиниць, визначити стиль і жанр, у яких працює письменник.

Виділення не вирішених раніше частин загальної проблеми. Хоча творчість Романа Іваничука вивчали багато дослідників, манера письма, індивідуальний стиль та почерк автора не було досліджено на достатньому рівні. Тобто, творчу діяльність письменника вивчали з літературознавчого погляду, але не досліджували мову автора. Відсутність таких досліджень зумовлена потребою дослідити багатогранну манеру письма митця, типові слова, діалекти, художні засоби.

Мета статті: за допомогою методів статистичного аналізу визначити особливості індивідуального стилю Романа Іваничука.

Виклад основного матеріалу. У таблицях нижче вказано інформацію про кількість букв, вжитих у словах та розподіл літер у розділах. Статистичні дані отримали за допомогою програми, яку написали на основі мови Python.

Як бачимо, найчастіше у повістях письменника трапляються слова, довжина яких 5 букв – 3807 слів (13,35%) та 4 букви – 3525 слів (12,36%). 2 букви – це довжина 3493 слів (12,25%). 1 букву мають 3268 слова (11,46%), 6 букв – 3256 (11,42%), 7 – 2713 (9,51%), 3 – 2245 (7,87%), 8 – 2243 (7,86%), 9 – 1698 (5,95%). Групу з найменшою довжиною становлять: 10 букв – 1049 слів (3,68%), 11 – 664 (2,33%), 12 – 330 (1,16%), 13 – 133 (0,47%), 14 – 49 (0,17%), 15 – 27 (0,09%), 16 – 14 (0,05%), 17 – 4 (0,01%), 21 – 1 (0,00%), 22 – 1 (0,00%), 23 – 1 (0,00%).

Отже, кількість значень (N) – 28521, максимальне значення (max) – 23, мінімальне значення (min) – 1, розмах (R) – 22, мода (Mo) – 5, медіана (Md) – 10,5, середнє (X) – 5,15, середнє квадратичне відхилення (σ) – 2,91, коефіцієнт варіації (v) – 0,5655, стандартна похибка (Sx) – 0,0172, відносна похибка (ε) – 0,0066.

Таблиця 1

Розподіл кількості букв у словах в повістях Романа Іваничука

Довжина слова	Кількість		Довжина слова	Кількість	
	Абсолютна	Відносна, %		Абсолютна	Відносна, %
1	3268	11,46%	11	664	2,33%
2	3493	12,25%	12	330	1,16%
3	2245	7,87%	13	133	0,47%
4	3525	12,36%	14	49	0,17%
5	3807	13,35%	15	27	0,09%
6	3256	11,42%	16	14	0,05%
7	2713	9,51%	17	4	0,01%
8	2243	7,86%	21	1	0,00%
9	1698	5,95%	22	1	0,00%
10	1049	3,68%	23	1	0,00%

Таблиця 2
Числові характеристики вибірки

Розділ твору	Весь текст
N	28521
Max	23
Min	1
R	22
Mo	5
Md	10,5
\bar{X}	5,15
σ	2,91
N	0,5655
S \bar{x}	0,0172
ε	0,0066

Таблиця 4
Числові характеристики вибірки

Розділ твору	Весь текст
N	5
Max	43896
Min	13588
R	30308
Mo	13588
Md	26143
\bar{X}	29361.8
σ	11224.05
N	0.3823
S \bar{x}	5019.547
ε	0.3351

Таблиця 3
Розподіл букв у розділах

Розділ	Кількість	
	Абсолютна	Відносна, %
I	13588	20.00%
II	22944	20.00%
III	26143	20.00%
IV	40238	20.00%
V	43896	20.00%

Отже, бачимо з таблиці, що у першому творі 13588 букв – (20,00%), у другому 22944 букв – (20,00%), у третьому 26143 букв – (20,00%), у четвертому 40238 букв – (20,00%), у п'ятому 43896 букв – (20,00%).

Бачимо з таблиці, що кількість значень (N) – 5, максимальне значення (max) – 43896, мінімальне значення (min) – 13588, розмах (R) – 30308, мода (Mo) – 13588.

Висновок. У проведеному дослідженні за допомогою методу статистичного аналізу було розглянуто твори, як от: «Соло на флейті», «Неспокутне», «Лісова повість», «Погоня за Пегасом», «І земля, і зело, і пісня». На основі даних з таблиць можна побачити, що найбільше у повістях письменника трапляються слова, довжина яких 5 букв – 3807 слів (13,35%) та 4 букви – 3525 слів (12,36%). Найбільше букв у 4 та 5 розділах: 40238 букв – (20,00%) та 43896 букв – (20,00%) відповідно.

Список літератури:

- Альтман Г. Мода та істина в лінгвістиці. Проблеми квантитативної лінгвістики. Чернівці : Рута, 2005. 109 с.
- Бук С. Основи статистичної лінгвістики : навчально-методичний посібник. Відп. ред. проф. Ф.С. Бацевич. Видавничий центр ЛНУ імені Івана Франка, 2008. 124 с.
- Васильев Л.М. Методы современной лингвистики. Уфа : Изд-во Башкирст, 1997. 182 с.
- Головин Б.Н. Язык и статистика: Просвещение, 1971. 189 с.
- Іваничук Р. І земля, і зело, і пісня: короткі повісті. Львів : Срібне слово, 2006. 160 с.
- Іваничук Р. Одна хлібина на двох: новели про любов. Вид. друге, без змін. Львів : Срібне слово, 2010. 160 с.
- Кульчицький І.М. Дослідження довжини речення та слова у творах Романа Іваничука. *Вісник Національного університету «Львівська політехніка»*. Серія : Інформаційні системи та мережі. Львів : Видавництво Львівської політехніки, 2017. № 872. С. 139–148.
- Нелюбин Л.Л. Толковый переводоведческий словарь : учеб. пособ., 3-е изд., перераб. Москва : Флинта, 2003. 320 с.
- Перебийніс В.С. Теоретичні та прикладні проблеми структурно-математичної лінгвістики. 1981. № 4.
- Хомицька І.Ю., Теслюк В.М. Метод статистичного аналізу англійських поетичних текстів. *Зб. наук.-техн. пр. : Науковий вісник НЛТУ України*. 2015. С. 350–356.

References:

- Altman, G. (2005). Moda ta istyna v lnhvistytsi/Problemy kvantytatyvnoyi lnhvistyky [Fashion and Truth in Linguistics. Problems of Quantitative Linguistics]. Chernivtsi: Ruta, 109 p.
- Buk, S. (2008). Osnovy statystychnoyi lnhvistyky [Fundamentals of Statistical Linguistics]: prof. F.S. Batsevich. Ivan Franko Publishing Center, 124 p.
- Vasilyev, L.M. (1997). Metody sovremennoy lingvistiki [Methods of Modern Linguistics]. Ufa: Bashkirst Publishing House, 182 p.
- Golovin, B.N. (1971). Yazyk i statistika [Language and statistics: Enlightenment], 189 p.
- Ivanychuk, R. (2006). I zemlya, i zelo, i pisnya: korotki povosti. [And the earth, and the green, and the song: short stories]. Lviv: Silver Word, 160 p.
- Ivanychuk, R. (2010). Odna khlibyna na dvokh: novely pro lyubov. [One loaf for two: stories about love]. View. second, unchanged. Lviv: Silver Word, 160 p.
- Kulchytskii, I.M. (2017). Doslidzhennya dovzhyny rechennya ta slova u tvorakh Romana Ivanychuka [Investigation of sentence length and words in the works of Roman Ivanychuk]. *Bulletin of the National University "Lviv Polytechnic". Series: Information Systems and Networks*. Lviv: Lviv Polytechnic Publishing House, no. 872, pp. 139–148.
- Neliubin, L.L. (2003). Tolkovyy perevodovedcheskiy slovar [Interpretive translation dictionary]: textbook. aids., 3rd ed., reworking. Moskva: Flint. 320 p.
- Perebyinis, V.S. (1981). Teoretychni ta prykladni problemy strukturno-matematychnoyi lnhvistyky [Theoretical and applied problems of structural and mathematical linguistics]. Vol. 4.
- Khomytska, I.Y., & Teslyuk, V.M. (2015). Metod statystychnoho analizu anhliys'kykh poetychnykh tekstiv [Method of statistical analysis of English poetic texts]. *Coll. scientific-technical BC: Scientific Bulletin of NLTU of Ukraine*, pp. 350–356.