

І. М. Кульчицький
 Національний університет “Львівська політехніка”,
 кафедра прикладної лінгвістики

ДОСЛІДЖЕННЯ ДОВЖИНИ РЕЧЕННЯ ТА СЛОВА У ТВОРАХ РОМАНА ІВАНИЧУКА

© Кульчицький І. М., 2017

Праця належить до одного із важливих напрямів квантитативних досліджень мови та мовлення – вивчення інформаційно-статистичних властивостей тексту. Розраховано довжину речення та слова у творах Р. Іваничука. Отримані результати зіставлено з аналогічними показниками в українській прозі.

Ключові слова: квантитативні дослідження, довжина речення, довжина слова, лінгвостатистика, стиль.

The article is dedicated to one of the most important areas of quantitative studies of language and speech that is the study of information and statistical properties of text. The length of sentences and words was calculated in the works of R. Ivanychuk. Received results were compared to the similar figures in Ukrainian prose.

Key words: quantative research, sentence length, word length, linguostatistics, style.

Вступ

Лінгвістична наука багата методами досліджень та підходами до тексту, проте останнім часом існує тенденція до використання квантитативних методів для потреб лінгвістики. Вважають, що найточніші результати можна отримати на стику цих двох наук. Фердинанд де Соссюр ще на початку ХХ століття звернувся до застосування статистичних методів для лінгвістичних досліджень [18].

Статистичні методи, які істотно видозмінюють та збільшують наші знання як про саму систему мови, так і про особливості її функціонування [8], використовують у мовознавстві у трьох умовних напрямках [20]:

- різноманітні кількісні підрахунки;
- побудова стохастичних лінгвістичних моделей з використанням отриманих у попередньому напрямі даних;
- перевірка статистичними методами різних гіпотез про ті чи інші лінгвальні явища.

Тому квантитативні дослідження української мови і сьогодні залишаються актуальними.

Постановка проблеми

Статистика – це математична наука, яка має на меті збирання, аналіз, тлумачення, пояснення та представлення даних. Однак цю науку використовують не лише для математичних досліджень. Сфера застосування статистики надзвичайно широка, починаючи від природничих та суспільних наук та закінчуючи бізнесом та політикою. Відомо, що статистику все частіше використовують для лінгвістичних досліджень, зокрема для вивчення особливостей авторської мови. Окрім цього, у мовознавстві використовують такі статистичні методи, як: “хі-квадратичний та варіативний аналіз, кластерний аналіз, аналіз головних компонентів та факторний аналіз” [10].

В. Перебийніс вважає, що всі складні системи підпорядковуються статистичним законам. Такими системами є живі організми, економіка, а також наука та людська мова. Науковець вказує на те, що необхідно використовувати статистичні методи у мовознавстві, адже мова є складною

системою, а відтак підпорядковується статистичним законам. Вона має кількісні характеристики, притаманні усім рівням системи. Зазначимо, що характеристики одиниць нижчого рівня істотно відрізняються від характеристик одиниць вищого рівня [15].

В. Перебийніс стверджує, що мова – це системи дискретних одиниць, а будь-які дискретні одиниці мають кількісні характеристики. В. Іванов зазначає: “...деякі основні кількісні характеристики мови, особливо істотні для лінгвіста, носять дуже простий характер, такою є зокрема розбіжність між числом слів (від десяти тисяч до ста тисяч), числом морфем (кілька тисяч), числом складів (від кількох сотень до кількох тисяч) і числом фонем (від 10 до 80). Згідно із висловленим вище припущенням ці співвідношення пов’язані зі структурою людської пам’яті” [12].

Явно чи неявно, але всі мовознавці визнають той факт, що у мові є кількісні характеристики. В. Перебийніс наводить приклади функціонування таких слів, як “більше”, “менше”, “багато” тощо. Такі оцінки не точні. Російський математик В. Буняковський іронічно зазначає: “...хай дозволено мені буде додати кілька слів про інше використання аналізу імовірностей...”, а також: “маючи подібні статистичні дані для двох чи кількох мов, можна порівнювати їх у різних відношеннях і результати набудуть певної авторитетності, яку на своє виправдання не завжди можуть забезпечити філологи за сучасного стану науки” [6].

Як система мова має свою структуру, тобто певний спосіб організації. Усі елементи мови організовані структурно. Це визначає їх взаємозв’язок між мовними елементами. Існують відношення і між одиницями одного рівня, і між одиницями різних рівнів. Кількість відношень є надзвичайно великою і постійно зростає. Наприклад, “якщо вимірювати відношення між фонемами у двофонемних утвореннях, то їхня кількість дорівнює кількості цих утворень, тобто вимірюється числом K^2 , де K – кількість фонем. У мові, в якій є 40 фонем, це число дорівнює $1600=40^2$ ” [15].

Існують пари фонем, у яких зв’язки між одиницями дорівнюють 0. Але навіть якщо одна третина цих зв’язків реалізується, кількість відношень залишається надто великою, щоб її рахувати. Якщо розглядати морфеми, склади чи словосполучення фонем, то їхня кількість також стрімко зростатиме. Науковець не може вирахувати дію кожного відношення, так само як і не може передбачити результат, він може лише припускати з певною імовірністю, пізнати яку йому допоможуть статистичні методи [15, с. 8].

Статистичні методи поширюються і на людське мовлення. Закони мови, закони сполучуваності одиниць, закони жанру, тема висловлювання впливають на мовлення людини. Всі ці фактори працюють в сукупності й виділити та визначити результат впливу окремого із них дуже важко. Але якщо сукупність цих факторів залишити незмінною, то наслідки будуть приблизно однакові, тобто мовлення характеризуватиметься певними рисами, які допоможе розкрити статистика. Також на будову мови впливає будова людського мозку, основними одиницями якої є об’єм пам’яті, швидкість запам’ятовування та відтворення того, що запам’яталось, характер фізіологічних реакцій організму. Ці фактори також діють сукупно та їхні характеристики вивчаються статистично [15, с. 8].

Робота над укладанням частотних словників також нерозривно пов’язана із підпорядкованістю мови й мовлення статистичним законам. У такому випадку лінгвісти працюють з великими обсягами матеріалу, маючи на меті встановити певні мовні характеристики. Їхня робота передбачає опрацювання матеріалу, неможливе без статистичних методів [15, с. 9].

Статистичні методи широко використовують у мовознавстві. Наприклад, у лексикографії статистику використовують для складання частотних словників, які допомагають підібрати найвживанішу лексику. Також статистичні методи пов’язані зі створенням методики навчання іноземної мови та оптимізації навчання [14, с. 211]. Статистичні методи відіграють важливу 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 уможливує встановлення семантичної близькості слів, визначає структуру семантичного поля, дослідження стилю письменника та атрибуції текстів, виявлення хронології мовних явищ [7, с. 66].

Отож, для досліджень у галузі лінгвістики використовують імовірно-статистичні методи математичної лінгвістики, а також методи та прийоми математичної теорії множин, а саме моделювання множин, визначення множин, взаємне накладання множин з метою їх порівняння. Порівняльні методи широко застосовують для вивчення тематичних та формальних класів мовних одиниць і, хоч кількісні методи, зазвичай, зараховують до практичної сфери, їх доволі широко використовують для вирішення теоретичних проблем мовознавства. Статистичні методи, які застосовують у лінгвістиці, не дуже відрізняються від застосовуваних в інших науках, наприклад, в експериментальній фонетиці використовують такий самий математичний апарат, що й у фізиці, у психолінгвістиці та соціолінгвістиці опрацьовують думку респондентів за допомогою тих самих методів конструювання шкал, що й у психології та соціології.

У лінгвістиці з'являються специфічні аспекти застосування кількісних методів, пов'язані з протиставленням мови та мовлення. Щодо системи мови кількісні методи застосовуються вкрай рідко, і лише коли йдеться про лексику (наприклад, кількісне вивчення етимологічного складу словника). Об'єктом застосування кількісних методів зазвичай є мовлення, точніше текст. Кількісні показники дають певну інформацію про текст. Той факт, що між мовними стилями та жанрами є певні відмінності, які мають статистичний характер, покладено в основу так званої статистичної стилістики.

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.

Аналогію залежності виявлено між конкретністю змісту та конкретністю мовлення. Наприклад, у текстах абстрактного значення, які представлені, передусім, науковими публікаціями, сумарний середній коефіцієнт абстрактності мовлення дорівнює 0,76 або 76 %, а конкретності 0,24 (24 %), для текстів конкретного змісту (наприклад художніх) також сумарно в середньому коефіцієнт абстрактності дорівнює 0,30 (30 %), конкретності – 0,70 (70 %). Кількісні методи використовують для класифікації текстів. Наприклад, для атрибуції текстів (визначення авторства), тому що більшість двосторонніх одиниць і конструкцій мови можуть слугувати основою для різних текстів або ж для їхньої кваліфікації. З теоретичної позиції використання статистичних методів у мовознавстві дає змогу доповнити структурну модель мови імовірнісними компонентами. Побудова такої моделі є завданням теоретичної лінгвістики, яке входить у компетенцію теорії мови [2, с. 40].

Аналіз досліджень та публікацій

Належить вважати, що початковою точкою появи в лінгвістиці кількісних методів, в сучасному розумінні, є ХХ ст., пов'язане із появою структурної лінгвістики. Цей напрям отримав назву завдяки посиленій увазі до структури мови, адже саме структура мови завжди була пріоритетною ділянкою для застосування кількісних критеріїв [3].

На горизонті ХІХ–ХХ ст. варто згадати й інших вчених, які також почали першими використовувати математичні, зокрема статистичні, методи для лінгвістичного опису. До них належать англійський вчений Дж. Юла, німецький науковець Е. Форстеманн, російські математики В. Буняковський, А. Марков та ін. Їх не хвилювало питання, чи результати їхніх досліджень будуть цікавити лінгвістів, а основним питанням для них було використання мовних елементів як ілюстративного матеріалу для побудови квантитативних методів чи для статистичних теорем [11, с. 115]. Науковець А. Росс [24] мав дещо протилежну думку, а саме він вважав, що теорія імовірностей і математична статистика можуть бути інструментом для перевірки і підтвердження тих лінгвістичних висновків, які допускають числове трактування. З вищевказаного можна зробити висновок, що між вченими того покоління не було однодумців щодо математичного напрямку в лінгвістиці [11, с. 115]. Також можна відзначити пізніші праці цього часу: праці з матлінгвістики: Р. Піотровський [16; 17] і В. Фукс [22], німецький вчений Дж. Ципф [26], дослідження про відношення між рангом і частотністю слова, які здійснили американці К. Шеннон та У. Уївер [25], дослідження теорії інформації для лінгвістичного опису Г. Хердана [23], який вивчав квантитативну лінгвістику.

Однак не було відкриттям те, що існував тісний зв'язок між звичайною і математичною мовами. У своїй праці А. В. Гладкий пише про те, що Л. Виготський у своїй книзі "Мислення і мова" писав: "Перший, хто побачив в математиці мислення, яке походить з мови, але перевершує її, був Декарт... Наша звичайна розмовна мова через притаманні їй коливання та граматичні й психологічні невідповідності знаходиться в стані рухомої рівноваги між ідеалами математичної і фантастичної гармонії, і в неперервному русі, який ми називаємо еволюцією" [9, 22].

На рубежі XIX–XX століття виникає системний підхід до науки. Чарльз Дарвін у своїй теорії еволюції органічного світу доводить, що органічний світ є єдиним цілим, яке складається з систематичних груп, пов'язаних певними відносинами. Д. І. Менделєєв досліджує періодичний закон хімічних елементів, на підставі якого всі сполуки об'єднуються у єдине ціле, що допомагає визначити те, що нам ще невідомо. Відтоді в науці, техніці та мистецтві ідея взаємозв'язків між всім, що існує, набуває все більшої популярності [1, 17].

Загальна теорія системи виникла в XX столітті. Вона не пов'язана з фізичною природою об'єктів і не намагається замінити собою інші науки, проте її метою є вивчення загальних питань структури системи та її управління. Теорія системи звертає увагу на цілісність об'єктів і зв'язки між ними, як у системі, так і з навколишнім світом.

У теорії інформації поняття навколишнього світу входить в поняття системи. Вважають, що множина утворює систему, якщо зв'язки певного типу між елементами однієї множини, тобто її внутрішні зв'язки, переважають зв'язки між елементами цієї множини і навколишнього світу, тобто зовнішні зв'язки. Першими мовознавцями, які підходили до описання мови і мовних елементів як системних зв'язків, були Фердинанд де Соссюр та Будуен де Куртене. Ф. Соссюр стверджував: "... Мова – це система, елементи якої утворюють ціле, а значимість одного елемента виникає тільки від одночасної наявності інших" [1, 18].

Ф. Соссюр розділяв та протиставляв два аспекти вивчення мови – синхронічний та діахронічний. На його думку, системні зв'язки в мові можна досліджувати лише у синхронії, оскільки час має властивість руйнувати будь-які системні зв'язки. Науковець також порівнював мову з грою в шахмати, у якій важливі правила та значення, а матеріал, з якого зроблені фігури, не має значення.

Пізніше науковці відмовлялися від поділу на синхронію та діахронію. Французький вчений Г. Гійом вважав, що мова є результатом нашого минулого і тому пропонував вивчати синхронію в розвитку. На його думку, зміни можуть відбуватися всередині системи без зміни механізму системи, так само вони можуть викликати зміни в самому механізмі. Такий підхід дає змогу не лише описати, а й пояснити мовні факти.

У сучасній лінгвістиці існує багато визначень системи мови. Вони зводяться до того, що мова – це семіотична функціональна система, яка має на меті збереження інформації в людському суспільстві та свідомості й обмін нею. Мова – система систем, всередині якої існує фонетична, морфологічна, лексична, синтаксична та інші підсистеми.

Також існують й інші погляди щодо визначення поняття мови. Наприклад, спеціалісти в галузі функціональної стилістики вважають, що в мові існує система функціональних стилів, підсистеми якої створені власне функціональними стилями, такими як науковий чи художній. Кожна із таких підсистем має специфічні особливості на лексичному, а деколи навіть на фонетичному та синтаксичному рівнях, що зумовлено специфікою спілкування у певній сфері людської діяльності. Однак варто зазначити, що така специфіка не утворює значних бар'єрів між стилями, тобто всі стилі певною мірою пов'язані між собою.

Представники теорії тексту вважають, що кожний окремий текст є системою, об'єднаною комунікативною цілісністю, логічними, граматичними та стилістичними зв'язками. Для лексикологів мова – це знакова система. Сучасна лексикологія комплексно розглядає лексичну систему, вказує на зв'язки у ній. Слово існує у всіх системах, притаманних мові як семіотичній системі: у синтаксисі, прагматиці та семантиці. Воно, наприклад, є не лише елементом лексико-семантичної системи мови, а й елементом всієї системи мови загалом, морфема є елементом системи словотворення, фонема – елементом звукової системи [1, 19].

Дослідження в цьому напрямі проводили [4; 5] такі зарубіжні вчені, як Габріель Альтман (Gabriel Altmann), Рейнгард Кюлер (Reinhard Köhler) (Німеччина); Петер Гжибек (Peter Grzybek), Еммеріх Келіх (Emmerich Kelih) (Австрія); Гейза Віммер (Geiza Wimmer) (Словаччина); Адам Павловскі (Adam Pawłowski), Ядвіґа Самбор (Jadwiga Sambor) (Польща); Юхан Тулдава (Естонія); Раймунд Піотровський, Анатолій Шайкевич (Росія) та інші. В Україні квантитативними дослідженнями мовних явищ займаються Володимир Широков [21], Максим Криґін [13], Валентина Перебийніс [19], Соломія Бук [5] та ін.

Виклад основного матеріалу

Для проведення досліджень твори письменника було поділено на дві групи – романи (“Черлене вино”, “Манускрипт з вулиці Руської”, “Вода з каменю”, “Саксаул у пісках”, “Хресна проща” та “Вогненні стовпи”) та новели (“Бузьків огонь”, “Батько”, “Злочин”, “Прут несе кригу”, “На порозі”, “Порвана фотокартка”, “Рододендри”, “Мить краси”, “Відплата”, “Тополина заметіль”, “Весільна”, “Побий мене!”, “За простибі”, “В дорозі”, “Чужий онук”, “Дім на горі”, “Тиша”, “Сповідь”, “Юра Фірман”, “Сива ніч”, “Настуня”, “Бузьки на Семеновій хаті”, “Зелений гомін”, “Трамвайна зупинка “Базар””, “Повернення”, “Любчина осінь”, “Срібна попільничка”, “Морська зірка”, “Мольфарська бартка”, “Цимбали”, “Знайдена фотографія”, “Саморобний пістоль”, “Люлька з червоного дерева”, “Намісто з карельської берези”, “У Косові над Гуком”, “Весілля в Теремі”, “Одна хлібина на двох”, “Тонконогий лошачок”, “Несподіване...”, “Байкал”, “За стіною туману”, “Відплиття на острів Цітеру”, “На перевалі”). Вибрані новели становлять більшу частину (43) новел письменника (51), а романи відображають творчість письменника у різні періоди. Довжину слів розраховано для романів та новел, а довжину речень — для новел.

Для кожної досліджуваної одиниці спочатку отримували абсолютні частоти її значень, з яких формували варіаційний ряд. Для одержаних варіаційних рядів обчислено середні та знайдено статистичні оцінки середньої частоти: середнє квадратичне відхилення, міра коливання середньої частоти, стандартну похибку відхилення та відносну похибку. Розрахунки виконували за стандартними формулами, поданими у [19].

Середнє обчислювали за формулою:

$$\bar{x} = \frac{\sum x_i n_i}{\sum n_i},$$

де x_i – варіанта; n_i – кількість появ варіант у досліді; i – номер варіанти.

Середнє квадратичне відхилення розраховували за формулою:

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2 n_i}{\sum n_i}},$$

де x_i – варіанта; n_i – кількість появ варіант у досліді; i – номер варіанти; \bar{x} – середнє значення.

Міру коливання середньої частоти визначали за формулою

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{\sum n_i}},$$

а стандартну похибку – за формулою:

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{\sum n_i - 1}},$$

де σ – середнє квадратичне відхилення; n_i – кількість появ варіант у досліді; i – номер варіанти; \bar{x} – середнє значення.

Відносну похибку розраховували за формулою:

$$\varepsilon = \frac{1,96 \sigma_{\bar{x}}}{\bar{x}},$$

де \bar{x} – середнє; $\sigma_{\bar{x}}$ – міра коливання середньої частоти.

Для підрахунку кількості фонем у словах було використано такі правила:

літери, які можуть позначати дві фонemi:

щ – /шч/ – у всіх позиціях;

ї – /ji/ – у всіх позиціях;

я – /ja/ – на початку слова, після голосного перед наступним приголосним, в кінці слова після голосного, після апострофа та м'якого знака, після й;
 я – /a/ – після м'якого чи напівм'якого приголосного, без апострофа;
 ю – /ju/ – на початку слова, після голосного перед наступним приголосним, в кінці слова після голосного, після апострофа та м'якого знака, після й;
 ю – /y/ – після м'якого чи напівм'якого приголосного, без апострофа;
 є – /je/ – на початку слова, після голосного перед наступним приголосним, в кінці слова після голосного, після апострофа та м'якого знака, після й;
 є – /e/ – після м'якого чи напівм'якого приголосного, без апострофа;
 інші варіації:
 дз – або дві фонемі /д/ і /з/, або одну /дз/;
 дж – або дві фонемі /д/ і /ж/, або одну /дж/;
 інші графічні одиниці:
 ь – позначає м'якість /'/
 ˘ – апостроф – після твердих приголосних перед я, ю, є, ї (які позначають два звуки), в кінці скороченого слова (переважно діалектний чи розмовний варіант).

Для проведення цього дослідження використано внесені у текст структурні позначки. Виділяючи речення, складні, підрядні та складені речення рахували як одне, а слова, написані через дефіс, як одне слово. Отримані результати було порівняно з аналогічними показниками у інших українських авторів.

Довжину слів, а точніше словоформ, вимірювали у фонемах. Для новел розподіл довжини слів подано у табл. 1.

Таблиця 1

Розподіл довжини слів у новелах Р. Іваничука

Довжина слова	Кількість		Довжина слова	Кількість	
	абсолютна	відносна, %		абсолютна	відносна, %
1	9047	10,110	12	906	1,010
2	11058	12,350	13	369	0,410
3	9449	10,560	14	155	0,170
4	12001	13,410	15	71	0,080
5	11199	12,510	16	28	0,030
6	10015	11,190	17	21	0,020
7	8667	9,680	18	12	0,010
8	6670	7,450	19	5	0,010
9	4895	5,470	20	3	0,004
10	3072	3,430	22	2	0,003
11	1861	2,080	25	1	0,002

Як бачимо, найчастіше у новелах письменника трапляються слова, довжина яких 4 – 12001 (13,41 %), 5 – 11199 (12,51 %) та 2 – 11058 (12,35 %) фонемі. Середню групу становлять слова, довжина яких 6 – 10015 (11,19 %), 3 – 9449 (10,56 %), 1 – 9047 (10,11 %), 7 – 8667 (9,68 %), 8 – 6670 (7,45 %) та 9 – 4895 (5,47 %) фонем. Слова, довжина яких понад 10 фонем, становлять тільки 3,8 % (10 – 3072 (3,43 %), 11 – 1861 (2,08 %), 12 – 906 (1,01 %), 13 – 369 (0,41 %), 14 – 155 (0,17 %), 15 – 71 (0,08 %), 16 – 28 (0,03 %), 17 – 21 (0,02 %), 18 – 12 (0,01 %), 19 – 5 (0,01 %), 20 – 3 (0,0 %), 22 – 2 (0,0 %), 25 – 1 (0,0 %)) усіх слів. Слова, довжина яких більша, ніж 19 фонем, –

складені, написані через дефіс, які ми рахували як одне слово. Статистичне опрацювання отриманих результатів дало такі показники: середнє значення – 5,06; середнє квадратичне відхилення – 2,83; міра коливання середньої частоти – 0,01; стандартна похибка відхилення – 0,01; відносна похибка – 0,02.

Розподіл довжини слів у романах Р. Іваничука подано у табл. 2. За зменшенням частоти довжини розмістились так: 4 – 61041 (12,63 %), 5 – 60338 (12,48 %), 6 – 54422 (11,26 %), 2 – 53196 (11,01 %), 1 – 49892 (10,32 %), 3 – 47655 (9,86 %), 7 – 46810 (9,68 %), 8 – 38261 (7,92 %), 9 – 29110 (6,02 %), 10 – 19841 (4,1 %), 11 – 11949 (2,47 %), 12 – 6119 (1,27 %), 13 – 2780 (0,58 %), 14 – 1028 (0,21 %), 15 – 468 (0,1 %), 16 – 201 (0,04 %), 17 – 114 (0,02 %), 18 – 53 (0,01 %), 19 – 32 (0,01 %), 20 – 30 (0,01 %), 21 – 11 (0,0 %), 22 – 7 (0,0 %), 24 – 3 (0,0 %), 23 – 2 (0,0 %), 25 – 2 (0,0 %), 26 – 1 (0,0 %).

Таблиця 2

Розподіл довжини слів у романах Р. Іваничука

Довжина слова	Кількість		Довжина слова	Кількість	
	абсолютна	відносна, %		абсолютна	відносна, %
1	9047	10,110	12	906	1,010
2	11058	12,350	13	369	0,410
3	9449	10,560	14	155	0,170
4	12001	13,410	15	71	0,080
5	11199	12,510	16	28	0,030
6	10015	11,190	17	21	0,020
7	8667	9,680	18	12	0,010
8	6670	7,450	19	5	0,010
9	4895	5,470	20	3	0,004
10	3072	3,430	22	2	0,003
11	1861	2,080	25	1	0,002

Як бачимо, у новелах за кількістю переважають слова з 2–4 фонем, а кількість слів з 5–7 фонемами приблизно однакова, що і показує табл. 3, у якій подано статистичні характеристики даних про довжину слів у романах та новелах.

Таблиця 3

**Статистичні характеристики довжини слів
у новелах та романах Р. Іваничука**

	Новели	Романи
Кількість різних довжин	22	26
Середнє значення	5,06	5,25
Середнє квадратичне відхилення	2,83	2,93
Міра коливання середньої частоти	0,01	0,00
Стандартна похибка відхилення	0,01	0,00
Відносна похибка	0,00	0,00

Порівняння отриманих результатів із середніми результатами стилів української мови [76, с. 178] показує, що вони повністю узгоджуються з відповідними показниками творів Р. Іваничука.

У табл. 4 показники романів та новел Р. Іваничука порівняно з аналогічними показниками інших українських прозаїків.

Таблиця 4

Довжина слова у творах українських прозаїків

	Середнє значення	Середнє квадратичне відхилення	Відносна похибка
А. Головка	4,74	0,10	0,034
О. Гончар	5,41	0,07	0,021
О. Довженко	4,73	0,08	0,027
П. Панч	5,28	0,29	0,091
М. Стельмах	5,30	0,16	0,048
Ю. Яновський	5,06	0,13	0,041
Новели Р. Іваничука	5,06	2,83	0,000
Романи Р. Іваничука	5,25	2,93	0,000

Як бачимо, середня довжина слів у новелах Р. Іваничука найближча до творів Ю. Яновського, а романів – до М. Стельмаха та П. Панча.

У табл. 5 подано розподіл довжини речень у новелах Романа Іваничука за кількістю слів (словоформ).

Таблиця 5

Довжина речень у новелах Р. Іваничука

Інтервал	Кількість		Інтервал	Кількість	
	абсолютна	відносна, %		абсолютна	відносна, %
1–3	389	7,82	34–36	112	2,25
4–6	823	16,54	37–39	83	1,67
7–9	774	15,55	40–42	76	1,53
10–12	523	10,51	43–45	58	1,17
13–15	403	8,10	46–48	61	1,23
16–18	308	6,19	49–51	45	0,90
19–21	271	5,45	52–54	38	0,76
22–24	228	4,58	55–57	35	0,70
25–27	210	4,22	58–60	28	0,56
28–30	148	2,97	> 60	239	4,80
31–33	124	2,49			

За зменшенням кількості отримані дані розмістились так: 4–6 – 823 (16,54 %), 7–9 – 774 (15,55 %), 10–12 – 523 (10,51 %), 13–15 – 403 (8,1 %), 1–3 – 389 (7,82 %), 16–18 – 308 (6,19 %), 19–21 – 271 (5,45 %), > 60 – 239 (4,8 %), 22–24 – 228 (4,58 %), 25–27 – 210 (4,22 %), 28–30 – 148

(2,97 %), 31–33 – 124 (2,49 %), 34–36 – 112 (2,25 %), 37–39 – 83 (1,67 %), 40–42 — 76 (1,53 %), 46–48 – 61 (1,23 %), 43–45 – 58 (1,17 %), 49–51 – 45 (0,9 %), 52–54 – 38 (0,76 %), 55–57 – 35 (0,7 %), 58–60 – 28 (0,56 %). Статистична обробка отриманих результатів дала такі показники: кількість різних довжин речень – 141, середнє значення – 19,57; середнє квадратичне відхилення – 22,46; міра коливання середньої частоти – 0,32; стандартна похибка відхилення – 0,32; відносна похибка – 0,03. Отримані дані порівняно з аналогічними показниками для творів різних стилів української мови (табл. 6) [19, с. 154].

Таблиця 6

**Довжина речень у творах різних стилів української мови
та у новелах Р. Іваничука**

	Драматургія, %	Художня проза, %	Поезія, %	Новели Р. Іваничука, %
1–3	49,73	8,78	11,74	7,82
4–6	29,07	18,60	18,87	16,54
7–9	12,14	18,65	23,02	15,55
10–12	5,20	16,01	18,33	10,51
13–15	1,86	12,17	8,79	8,10
16–18	1,10	7,83	7,26	6,19
19–21	0,44	5,19	5,06	5,45
22–24	0,28	3,74	2,45	4,58
25–27	0,08	2,79	1,70	4,22
28–30	0,02	1,64	0,95	2,97
31–33	0,02	1,50	0,46	2,49
34–36	0,04	0,80	0,50	2,25
37–39	0,00	0,75	0,17	1,67
40–42	0,00	0,25	0,12	1,53
43–45	0,00	0,25	0,12	1,17
46–48	0,00	0,25	0,12	1,23
49–51	0,02	0,20	0,04	0,90
52–54	0,00	0,20	0,00	0,76
55–57	0,00	0,10	0,12	0,70
58–60	0,00	0,05	0,00	0,56
> 60	0,00	0,50	0,17	4,80

Аналіз даних показує, що отримані у дослідженні результати відповідають даним для прози української мови. Єдина й істотна відмінність – великий відсоток довгих речень, що дає підстави виокремити цю ознаку як характерну для новелістики Р. Іваничука.

Висновки

Проведені дослідження показали, що за окремими статистичними характеристиками твори Р. Іваничука відповідають показникам прозових творів всієї української літератури. Виявлення квантитативних характерних особливостей творів письменника потребує додаткових досліджень інших його творів та розширення кількості показників.

1. Арнольд И. В. *Основы научных исследований в лингвистике: учеб. пособ.* – М.: Высш. шк., 1991. – 140 с.
2. Баранов А. Н. *Введение в прикладную лингвистику: учеб. пособ.* / А. Н. Баранов. – М.: Эдиториал УРСС, 2001. – 360 с.
3. БСЭ – Большая советская энциклопедия [Электронный ресурс]. – 1978. – Режим доступа : dic.academic.njournals.nsfibse/
4. Бук С. *Сучасні методи дослідження мови письменника у слов'язнознавстві* / С. Бук // *Проблеми слов'язнознавства.* – 2012. – Вип. 61. – С. 86–95.
5. Бук С. *Лінгвостатистичний опис “Не спитавши броду” Івана Франка* / Соломія Бук [Електронний ресурс] – Режим доступу : http://www.lnu.edu.ua/faculty/Philol/www/visnyk/55_2011/55_2011_Vuk.pdf
6. Буняковский В. Я. *О возможности введения определительных мер доверия к результатам некоторых наук наблюдательных и преимущественно статистики* / В. Я. Буняковский // *Современник* – 1847. – Т. 3, раздел 2. – С. 48.
7. Васильев Л. М. *Методы современной лингвистики* / Л. М. Васильев. — Уфа : Изд-во Башкирст. ун-та, 1997. – 182 с.
8. Верховин С. С. *О статусе количественных методов в лингвистике* / С. С. Верховин [Електронний ресурс]. – Режим доступу : <http://cyberleninka.ru/article/n/o-statuse-kolichestvennyh-metodov-v-lingvistike>
9. Гладкий А. В. *Математические методы изучения естественных языков* / А. В. Гладкий // *Математическая логика, теория алгоритмов и теория множеств : сб. работ*, 1973. – С. 95–108.
10. Головин Б. Н. *Язык и статистика* / Б. Н. Головин. – М.: Просвещение, 1971. – 189 с.
11. Звегинцев В. А. *Очерки по общему языкознанию* / В. А. Звегинцев. – М.: МГУ, 1962. – 382 с.
12. Иванов В. В. *Некоторые проблемы современной лингвистики* / В. В. Иванов // *Народы Азии и Африки.* – 1963. – № 4.
13. Кригін М. Ю. *Дослідження інформаційно-статистичних властивостей українського тексту* / М. Ю. Кригін, В. А. Широков // *Математические машины и системы.* – ИПМС НАНУ, 2000. – № 1. – С. 120–127.
14. Нелюбин Л. Л. *Толковый переводоведческий словарь : учеб. пособ.* / Л. Л. Нелюбин. – 3-е изд., перераб. – М.: Флинта, 2003. – 320 с.
15. Перебийніс В. С. *Кількісні та якісні характеристики системи фонем сучасної української літературної мови* / В. С. Перебийніс. – К.: Наукова думка, 1970.
16. Пиотровский Р. Г. *Инженерная лингвистика и теория языка [Текст]* / Р. Г. Пиотровский. – Л.: Наука, С 1979. – 111 с.
17. Пиотровский Р. Г. *Математическая лингвистика: учеб. пособ.* / Р. Г. Пиотровский. – М. : Высш. шк., 1977. – 383 с.
18. Сосюр Фердінан де. *Курс загальної лінгвістики* / Фердінан де Сосюр: пер. з фр. А. Корнійчук, К. Тищенко. – К.: Основи, 1998. – 324 с.
19. *Статистичні параметри стилів / за ред. В. С. Перебийніс.* – К.: Наукова думка, 1967. – 260 с.
20. Фрумкина Р. М. *Роль статистических методов в современных лингвистических исследованиях* // Шаумян С. К. *Математическая лингвистика* / отв. ред. С. К. Шаумян. – М.: Наука, 1973. – С. 156–183.
21. Широков В. А. *Інформаційна теорія лексикографічних систем* / В. А. Широков. – К.: Довіра, 1998. – 331 с.
22. Fucks W. *Mathematische Analyse von Sprachelementen, Sprachstil und Sprachen* / W. Fucks. – Koln und Oplade, 1955. – 110 p.
23. Herdan G. *Language as Choice and Chance* / G. Herdan. – Noordhoff : Groningen, 1956. – 356 p.
24. Ross A. *Philological Probability Problems* / I. A. Ross. – Statist. Soc. – Vol. XII. – 1950. – P. 19–59.
25. Shannon C., Weaver W. *The Mathematical Theory of Communication* / C. Shannon, W. Weaver. – Univ of Illinois Press, 1949. – 117 p.
26. Zipf G. K. *Human Behavior and the Principle of Least Effort* / G. K. Zipf. – Addison-Wesley, 1949. – 573 p.