

Міністерство освіти і науки України
Чернівецький національний університет імені Юрія Федьковича

КАЛИМОН ЮЛІЯ ОЛЕКСАНДРІВНА 

УДК 811.161.2'374.2:821.161.2-32[Стеф]; 004](043.5)

**СТРУКТУРНО-ІНФОРМАЦІЙНА МОДЕЛЬ СЛОВНИКА МОВИ
НОВЕЛ ВАСИЛЯ СТЕФАНІКА**

10.02.01 — українська мова

Автореферат дисертації на здобуття наукового ступеня
кандидата філологічних наук

Чернівці – 2020

Дисертація є рукописом.

Роботу виконано на кафедрі прикладної лінгвістики Національного університету «Львівська політехніка», Міністерство освіти і науки України

Науковий керівник –

кандидат технічних наук, доцент
КУЛЬЧИЦЬКИЙ Ігор Маркіянович,
Національний університет
«Львівська політехніка»,
доцент кафедри прикладної лінгвістики.

Офіційні опоненти:

доктор філологічних наук, доцент
СИТАР Ганна Василівна,
Донецький національний університет
імені В. Стуса, професор кафедри
загального та прикладного мовознавства і
слов'янської філології;

кандидат філологічних наук, доцент
БЕРКЕЩУК Інна Степанівна,
Кам'янець-Подільський національний
університет імені Івана Огієнка,
доцент кафедри української мови.

Захист відбудеться „ 7 ” грудня 2020 р. о 10 годині на засіданні спеціалізованої вченої ради Д 76.051.07 Чернівецького національного університету імені Юрія Федьковича (58001, м. Чернівці, вул. Садова, 5).

Із дисертацією можна ознайомитися в Науковій бібліотеці Чернівецького національного університету імені Юрія Федьковича (58002, м. Чернівці, вул. Лесі Українки, 23) та на офіційному сайті Чернівецького національного університету імені Юрія Федьковича (<http://specrada.chnu.edu.ua/index.php?page=ua/07/baza>).

Автореферат розіслано „ 6 ” листопада 2020 р.

Учений секретар
спеціалізованої вченої ради



А. А. Король

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

У дисертації розпрацьовано структурно-інформаційну модель словника мови новел Василя Стефаника, що ґрунтується на корпусі текстів останніх прижиттєвих видань творів визнаного майстра експресіоністичної новели та відповідає основним текстологічним засадам укладання словників мови письменників.

Мова художніх творів як сукупність мовних засобів, які вирізняють індивідуальність митця, упродовж тривалого часу є об'єктом широкого зацікавлення та детального лінгвістичного опису. Цьому сприяє антропоцентричність та багатовекторність наукової парадигми сучасної лінгвістики.

Теоретичні та практичні проблеми лексикографічної діяльності на матеріалі різних мов розглядали Ю. Апресян, В. Виноградов, О. Демська, В. Дубічинський, Л. Зґуста, Ю. Караулов, Н. Клименко, С. Лендау, В. Перебийніс, Л. Полюга, В. Широков, та ін. До проблем *авторської лексикографії* зверталися С. Бук, В. Ващенко, О. Карпова, І. Ковалик, М. Лесюк, Л. Шестакова та ін.; *лінгвістики тексту та художнього дискурсу* – І. Бехта, А. Загнітко, Н. Кондратенко, І. Кочан, О. Селіванова та ін.; *ідіолекту письменника* – Н. Гуйванюк, М. Крупа, Л. Марчук, А. Мойсієнко, Н. Сологуб, Г. Ситар, Л. Ставицька та ін.; *концептуальної та мовної картин світу письменника* – І. Беркещук, В. Виноградов, В. Жайворонок, О. Потєбня, С. Шабат-Савка; *діалектної лексики в художніх творах українських письменників* – В. Грещук, О. Кульбабська, І. Матвіяс, Н. Руснак, М. С. Скаб і М. В. Скаб; *статистичного опрацювання корпусів текстів* – С. Бук, С. Еверт, І. Кульчицький, В. Левицький, В. Перебийніс, Г. Ситар, М. Стабз та ін. Корпуснообазовані та статистичні методи у галузі лексикографії дають змогу якнайповніше осягнути та проаналізувати систему мови, у синхронійному та діахронійному зрізах. Утім, досі є лакуни в царині лексикографічного опрацювання західноукраїнського варіанта української літературної мови на зламі XIX та XX століть.

Актуальність дослідження зумовлена тим, що воно знаходиться у руслі сучасних досліджень, спрямованих на комплексне вивчення функцій мовних одиниць різних рівнів у художньому тексті, а також нарощенням кількості корпуснообазованих та статистичних досліджень особливостей західноукраїнського варіанта української літературної мови на зламі XIX та XX століть, які представлено у прозових творах письменників того періоду.

Зв'язок роботи з науковими програмами, планами, темами. Дисертацію виконано в межах однієї з наукових тем кафедри прикладної лінгвістики Національного університету „Львівська політехніка”, а саме

„Застосування новітніх технологій для оптимізації інформаційних процесів природною мовою”, затвердженої Міністерством освіти і науки України (номер державної реєстрації 0114U001239). Тему дисертації затверджено на засіданні Вченої ради Національного університету „Львівська політехніка” (протокол № 61 від „23” квітня 2013 р.), уточнено на засіданні Вченої ради Національного університету „Львівська політехніка” (протокол № 60 від „23” грудня 2019 р.), та на засіданні бюро Наукової ради „Закономірності розвитку мов та мовної діяльності” НАН України (протокол № 1 від „11” лютого 2020 року).

Мета дослідження – розробити структурно-інформаційну модель словника мови новел Василя Стефаника на матеріалі корпусу останніх прижиттєвих видань та апробувати часткові статистичні обчислення, які виформовують статистичний профіль автора на прикладі його творів.

Досягнення мети передбачає виконання таких **завдань дослідження**:

- систематизувати та синтезувати теоретико-критичні здобутки міждисциплінарного підходу до укладання авторських словників;

- визначити термінологічний обсяг терміна „ідіолект письменника” у мовознавчих студіях і розглянути його як систему різнорівневих мовних одиниць;

- напрацювати методологічний алгоритм побудови корпусу творів Василя Стефаника;

- визначити та проаналізувати статистичні відомості, отримані з корпусу;

- порівняти отримані статистичні відомості з аналогічними показниками української художньої прози загалом та окремих письменників, а також описати притаманні риси ідіолекту Василя Стефаника.

Об’єктом дослідження є лексика новел Василя Стефаника.

Предметом дослідження є різнорівневі структурні та кількісно-дистрибутивні характеристики мовних засобів новел Василя Стефаника, що визначають особливості ідіолекту письменника.

Матеріалом дослідження слугував упорядкований електронний текстовий масив – корпус – із 57 новел, виданих за життя Василя Стефаника, загальним обсягом 53978 слововживань (12834 слів та словоформ).

Джерельна база дослідження – прижиттєві опубліковані твори Василя Стефаника. Також джерелом дослідження слугували словники: „Словник української мови: в 11 тт.” (1970–1980), „Словник української мови Online. Томи 1-10 (А-ОБМІЛЬ)”, „Частотний словник сучасної

української художньої прози: у 2 т.” (1981), „Гуцульські говірки: Короткий словник” (1997), „Матеріали до словника гуцульських говірок Косівська Поляна і Росішка Рахівського району Закарпатської області” (2005), „Скарби гуцульського говору: Березови” (2008), „Словник буковинських говірок” (2005), „Словник українських говірок Карпатського регіону: пояснення та походження слів” (2012), „Тлумачний словник-довідник гуцульських говірок” (2016).

Методика дослідження зумовлена метою та завданнями праці та лежить в площині висвітлення ідіолектних аспектів літературно-художнього прозового тексту, зорієнтованому на студіях функцій лексики як складника ідіолекту Василя Стефаника. У праці поетапно застосовано загальнонаукові та лінгвістичні методи. На першому етапі задіяно *аналіз і синтез* – для відбору релевантного теоретичного матеріалу, а на наступних – *індуктивно-дедуктивний метод* – для спостереження, опису та класифікації різнорівневих мовних одиниць. Для опису ідіолектних аспектів літературно-художнього прозового тексту Василя Стефаника використано *дескриптивний метод*. *Метод моделювання* використано для розуміння того, як структурно-інформаційно влаштовано словник мови письменника, взаємозв'язки його елементів, основні властивості, етапи його укладання та місце в загальній типології словників. *Корпуснобазований метод* задіяний для створення корпусу прозових текстів новел Василя Стефаника з дотриманням основоположних засад укладання лінгвістичних корпусів; *структурний* – для аналізу ідіолектних засобів на тлі лексичного та морфологічного рівнів мови; *метод словникових дефініцій* – для з'ясування структури значення лексичних одиниць; *контекстний метод* – для опису смислів, актуалізованих у текстовій канві художнього твору; *статистичний метод* – для підрахунку частоти вживання різнорівневих мовних одиниць; *кількісний аналіз* – з метою встановлення повторюваності, домінування, співвідношення певних мовних одиниць у новелах Василя Стефаника, які виформовують характерні риси ідіолекту письменника, *порівняльний аналіз* дав змогу виявити спільні та відмінні риси, притаманні Василю Стефанику та художній українській прозі загалом чи окремим письменникам. Застосовані методи та прийоми вможливають визначення того, яке значення лексичної одиниці має увійти у словник мови новел Василя Стефаника, та надати різноманітні відомості про мову новел для укладання різнопланових лексикографічних праць на основі прижиттєвих опублікованих новел автора (словопокажчика, конкордансу, частотного словника, словника мови новел Василя Стефаника).

Наукова новизна дослідження полягає в тому, що в дисертації *вперше* створено структурно та морфологічно роззначений корпус текстів

новел Василя Стефаника, виданих за життя письменника. *Уперше* запропоновано структурно-інформаційну модель словника мови новел Василя Стефаника та їх відповідне наповнення. У дослідженні *вперше* кількісно проаналізовано частиномовний, загальномовний / діалектний, авторський та персональний розподіл лексики; вираховано індекс епітетизації, дієслівних означень і номінальності. Запропоновано тематичний розподіл діалектної іменникової лексики та виокремлено провідні її концепти відносно мовної картини світу (за В. Жайворонком). *Уперше* визначено коефіцієнт винятковості лексики новел; здійснено порівняльний аналіз уживання назв кольорів та порівняльних іменних конструкцій із словом *як* у новелах та творах сучасників письменника; на прикладі найчастотніших лексем апробовано колокаційний підхід із застосуванням показників асоціації (що визначають наскільки сильним є зв'язок між словами), а саме *t-критерій* (*T-score*) та *взаємної інформації* (*Mutual Information*) задля вияву провідного значення багатозначного слова. Завдяки здійсненим *уперше* статистичним дослідженням різнорівневих одиниць мови та тексту (символів, біграм, слів та словоформ, речень), доведено близькість новел, виданих за життя Василя Стефаника, до показників української сучасної художньої прози загалом та окремих письменників.

Практичне значення отриманих результатів полягає в тому, що в дисертації доповнено і розвинуто наукові уявлення про вивчення ідіолекту письменника з застосуванням корпуснобазованих технологій і статистичного апарату. Результати, одержані в ході дослідження, можуть бути використані в лекціях та практичних заняттях із навчальних дисциплін „Діалектологія української мови” (розділ „Південно-західне наріччя. Територіальні діалекти”), „Сучасна українська мова. Лексикологія” (розділ „Аспекти й методи дослідження в лексикології”), „Стилістика” (розділ „Художній стиль української мови”), „Корпусна лінгвістика” (розділ „Корпусні інструменти для лінгвістичних досліджень”), спецкурсів із вивчення ідіолекту письменника, лінгвістичного аналізу тексту, для лінгвістичних наукових досліджень та виконання студентських кваліфікаційних наукових робіт.

Апробація матеріалів дисертації. Основні результати дисертаційного дослідження представлено та обговорено на засіданнях кафедри прикладної лінгвістики Національного університету „Львівська політехніка” (2013-2020), а також на *п'ятнадцятиох* міжнародних наукових конференціях конференціях: „Пріоритети германського і романського мовознавства” (Луцьк, 2012, 2013, 2014, 2015, 2017, 2018, 2019); „Мови і світ: дослідження та викладання” (Кіровоград, 2013); „Людина. Комп'ютер. Комунікація” (Львів, 2013, 2017, 2019);

„Комп’ютерні науки та інформаційні технології” CSIT (Львів, 2013); „Горизонти прикладної лінгвістики і лінгвістичних технологій : MegaLing’2012” (Київ, 2012); „Сучасна україністика : проблеми мови, літератури та культури” (Оломоуць, Чеська Республіка, 2018); одній Всеукраїнській науковій конференції „Актуальні тенденції сучасного гуманітарного дискурсу” (Дрогобич, 2019) та одному міжнародному науково-теоретичному та навчально-практичному семінарі „Лінгвіст-програміст” (Вінниця, 2018).

Публікації. Основні положення дисертації викладено у *сімнадцяти* публікаціях (із них – *дев’ять* одноосібних та *вісім* статей у співавторстві), серед яких *десять* статей надруковано у наукових фахових виданнях України (4,31 друк. арк.), *одну* – у періодичному виданні іншої держави (0,39 друк. арк.); *три* – у матеріалах міжнародних наукових конференцій і семінарів (0,77 друк. арк.) та *три* праці (1,13 друк. арк.), що додатково відображають наукові результати дисертації. Загальний обсяг публікацій становить 6,60 друк. арк.

Особистий внесок здобувача в публікаціях у співавторстві.

Серед наукових праць, опублікованих у співавторстві з І. Кульчицьким та І. Ліхнякевичем, дисертанткою прослідковано етапи становлення, окреслено основні об’єкти та проблеми застосування методу ентропії біграм задля визначення інваріанту твору („Ентропія біграм як метод формального аналізу художніх творів”); наведено огляд відмінностей термінологічного означення стилю мови письменника, основні причини ототожнення цих термінів, окреслено принципи відмінності їх тлумачення („Ідіолект, ідіостиль, індивідуальний стиль. Тотожне чи різне?”); проаналізовано найпоширеніші типи лексикографічного опису мови письменників („Лексикографічні аспекти опису мови письменників”); зібрано та опрацьовано фактичний матеріал стосовно розвитку корпусів національних мов („Деякі аспекти використання лінгвістичних корпусів словянських мов”); проаналізовано основні здобутки та сфери застосування паралельних корпусів для досліджень у галузі контрастивної лінгвістики, лексикографії („Деякі аспекти створення та використання паралельних корпусів”). Серед наукових праць, опублікованих у співавторстві з І. Кульчицьким, дисертанткою здійснено порівняльний аналіз частин мови деяких новел та їх відповідних перекладів англійською мовою, виформувано висновки дослідження („Кількісний аналіз новел В. Стефаніка та їх перекладів англійською мовою”); здійснено огляд основних напрямків дослідження творчості Василя Стефаніка, окреслено основні етапи створення корпусу новел („Деякі аспекти створення корпусу художніх творів В. С. Стефаніка”).

Структура та обсяг дисертації. Дисертаційне дослідження складається з анотацій двома мовами, списку опублікованих праць автора, вступу, трьох розділів, висновків, списку використаних джерел із 404 позицій (із них 14 позицій джерел довідкового матеріалу, 6 позицій джерел ілюстративного матеріалу), 16 додатків загальним обсягом 93 сторінки (з них 2 діаграми, 2 фрагменти словникових статей, 5 рисунків, 72 таблиці, списку опублікованих праць за темою дисертації). Загальний обсяг роботи становить 312 сторінок, із них – 162 сторінки основного тексту.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** висвітлено актуальність дослідження, зв'язок роботи з науковими програмами, планами, темами, сформульовано мету та завдання, представлено об'єкт і предмет, описано матеріал та джерельну базу, методологічний інструментарій дослідження, визначено практичне значення і наукову новизну дисертації, а також зазначено інформацію про кількість публікацій із теми дослідження й апробацію результатів на наукових конференціях.

У **першому розділі** „**Теоретичні аспекти створення структурно-інформаційних моделей словників у сучасній лексикографії**” з'ясовано стан сучасної української лексикографії, її основні здобутки та перспективи (п. 1.1), визначено поняття „**моделі**” як вищого рівня наукового пізнання, окреслено сутнісні властивості й ознаки моделі (п. 1.2), розроблено структурно-інформаційну типологію словників (п. 1.3), проаналізовано й описано структурно-інформаційне наповнення наявних словників мови письменників різних типів.

Стрижневою проблемою цієї праці є мова новел Василя Стефаника як приклад західноукраїнського варіанта української літературної мови на зламі XIX та XX століть і її лексикографічне опрацювання із застосуванням корпусного та статистичного інструментарію, що корелює з сучасними тенденціями світового словникарства (Н. Дарчук, Ю. Караулов, В. Перебийніс, В. Широков).

Побудова моделі є найвищим рівнем абстракції в науковому пізнанні та полягає в узагальненні основних властивостей досліджуваного об'єкта (І. Кульчицький, Я. Неуймін, В. Штофф). Модель словника письменника має такі ознаки: 1) за формою подання – інформаційна (структурно-змістова за предметом моделювання; знакова за засобами моделювання); 2) за галуззю застосування – навчальна та дослідна; 3) за галуззю знань – лінгвістична; 4) за формою наповнення – корпуснобазована; 5) за часовим чинником – статична.

Структурно-інформаційна модель словника, незалежно від типу, містить три рівні: макро-, мікро- та медіорівень (Е. Агрікола). До рівня макроструктури словника належать сукупність словникових статей, а також додаткові елементи (джерела словника, передмова, опис структури, список скорочень тощо). Відтак на макроструктурному рівні до реєстру словника мови новел Василя Стефаника увійдуть усі повнозначні (лематизовані) та службові частини мови. До питання організації мікроструктури, як рівня, який відбиває структурно-інформаційну будову словникової статті, належить підходити насамперед із міркувань практичного застосування словника (О. Демська, В. Дубічинський, С. Лендау). Відтак будова словникової статті має усталену структуру, яка лінійно представлена двома обов'язковими елементами: лівобічною (реєстрове слово) та правобічною (інформація про реєстрову одиницю) частиною (І. Кульчицький). Аналіз наявних авторських словників уможливив опис таких основних складників словникової статті: реєстрове слово з відміткою про частоту вживання у творах, граматичний / стилістичний коментар, тлумачення, цитування, словотвірна парадигма та фразеологічне вживання (де потрібно) із відміткою про частоту вживання (С. Бук, В. Ващенко, І. Ковалик, М. Лесюк та ін.). Зона тлумачення традиційно вимагає особливого підходу, зважаючи на поняття „лексичне значення слова” (Л. Лисиченко, Ж. Соколовська).

Найпоширеніші типи тлумачень у словниках мови письменника: констатувальні (фіксують лише те, як слово вживають і розуміють), вербальні (характеризують слово описово), номінальні (значення слова подають через значення іншого слова) / реальні (розкриття, опис значення слова, його сутності), предикативні (мають форму речень). Серед способів подання провідними є контактні (опис значення засобами звичайної мови, у вигляді деякого тексту, з використанням достатньої термінології певної сфери дійсності), здебільшого аналітичні (через вказівку на широкий клас і особливі ознаки; шляхом перерахування предметів, дій, ознак тощо, які позначають цим словом; за допомогою вказівки на не пов'язані ієрархічними відношеннями ознаки (з використанням слів *те, що, щось* тощо), рідше – відсылні (відсилання до тлумачення іншого слова: наприклад, *те саме, що*).

До найпоширеніших типів словників мови письменників належать: словопоказчики; конкорданси; частотні; глосарії, а також диференційні словники. Усе актуальнішим стає розроблення словника мови письменника ідеографічного типу (С. Бук, О. Карпова, Л. Шестакова).

У другому розділі „**Парадигма методів вивчення словника творів письменника**” окреслено теоретико-методологічні підходи до вивчення ідіолекту письменника як представника певного етносу на

певному часовому зрізі, а також термінологічний обшир понять на позначення сукупності мовно-виражальних засобів письменника (п. 2.1), описано особливості квантитативного, колокаційного та корпуснобазованого підходів до вивчення мови творів письменників (п. 2.2 та 2.3); описано принципи організації вибірки й укладання корпусу текстів новел Василя Стефаника (пп. 2.2.1, 2.2.2); схарактеризовано методологічну основу дослідження й окреслено п'ять етапів студіювання мови новел Василя Стефаника (пп. 2.1.1).

Індивідуально-авторський стиль формується під впливом взаємозумовлених факторів: мовних і позамовних, соціальних і політичних, філософських і світоглядних, і є виявом мовної особистості автора, його свідомості та мовної картини світу (І. Беркешук, Ю. Караулов, Л. Марчук, Г. Ситар). Важливим постає питання розмежування термінів на позначення мовних засобів майстрів слова, а саме понять ідіолекту, ідіостилю та індивідуального стилю (Н. Сологуб, Л. Ставицька). У цьому дослідженні ми дійшли висновку, що проміжне місце мови творів письменника на перетині двох наук – мовознавства й літературознавства – зумовлює неусталеність цих понять. Услід за Н. Сологуб та Л. Ставицькою, *ідіолект* розуміємо як сукупність мовних засобів, які не є сталою величиною, а тим, на що впливає й особистість і світогляд самого автора.

Створення корпусу текстів творів певного письменника першочергово полягає у виборі основного тексту творів, які увійдуть до нього, їх переведенні в електронну форму та нормалізації. Відтак ідуть етапи роззначення (від англ. – annotation) та лематизації (різного ступеня деталізації), які зазвичай здійснюють напівавтоматично (С. Бук, Н. Дарчук, І. Кульчицький, Г. Ситар).

Кількісні характеристики, які отримують із текстів, належать до таких лінгвістичних явищ, на які сам автор не звертає особливої уваги: вони властиві його текстам, але не залежать від його свідомості (В. Левицький, В. Перебийніс). Серед основних параметрів художнього тексту для встановлення індивідуальних рис стилю письменника виокремлюють: розподіл слів за частинами мови в реченнях, загальний розподіл довжини слова, загальний розподіл довжини речень, лексичний спектр текстів на рівні словника, середня довжина речень і середня довжина слова. Найінформативнішим вважають аналіз виокремлених параметрів текстів порівняно з іншими письменниками або із загальномовними показниками (С. Бук, Н. Єсипенко). Традиційні методи вивчення синтагматичних відношень між словами на сьогодні доповнені корпусно-статистичними, які засновано на аналізі кількісних характеристик ступеня сполучуваності слів у певному тексті (С. Еверт, Г. Ситар, М. Стабз). Такий підхід, на нашу думку, сприяє виокремленню провідного значення полісемантичного

реєстрового слова та дає змогу побудувати зону тлумачення словникової статті мови творів письменника за принципом актуалізованого значення цього слова.

Багатоплановість антропоцентричної парадигми сучасної лінгвістики та підходів, задіяних у нашому дослідженні, зумовлюють вибір релевантних методів дослідження: реферативний аналіз використано для опрацювання праць із теорії та практики словникарства на сучасному етапі й у розрізі авторської лексикографії, а також для вивчення підходів до трактування мовно-виражальних засобів письменників у сучасній лінгвістиці. За допомогою аналізу та синтезу інформації виявлено особливості структурно-інформаційної будови словників мови письменників, зокрема інтегрального типу, сформульовано висновки дослідження. Метод моделювання задіяно для розуміння того, як структурно-інформаційно влаштовано словник мови письменника, взаємозв'язки його елементів, основні властивості, правила його укладання та місце в загальній типології словників. Корпуснобазований метод застосовано під час створення корпусу текстів новел Василя Стефаника із дотриманням основоположних засад, які висувають до лінгвістичних корпусів; структурний метод – для аналізу ідіолектних засобів на тлі лексичного та морфологічного рівнів мови; методика компонентного аналізу із залученням лексикографічного методу – для з'ясування структури значення лексичних одиниць; контекстний метод – для опису смислів, актуалізованих у текстовій канві художнього твору; статистичний метод – для підрахунку частоти вживання різнорівневих мовних одиниць; кількісний аналіз – для встановлення повторюваності, домінування, співвідношення певних мовних одиниць у текстах новел Василя Стефаника, які становлять характерні риси ідіолекту письменника. Порівняльний аналіз дав змогу виявити спільні та відмінні риси, притаманні мові Василя Стефаника та художній українській прозі загалом чи деяким іншим письменникам зокрема.

Методологія дослідження структурно-інформаційної моделі словника мови новел Василя Стефаника передбачає кілька етапів.

На **першому етапі** проаналізовано й узагальнено теоретичні засади авторської лексикографії для чіткого розуміння проблеми та перспектив власного дослідження. Проаналізовано й узагальнено теоретичні засади авторської лексикографії та наявні приклади словників мови письменників.

На **другому етапі** здійснено відбір емпіричного матеріалу для створення дослідницького корпусу новел Василя Стефаника з дотриманням засадничих принципів укладання лінгвістичних корпусів, а саме: однорідності, повноти, роззначення для подальшого аналізу фактів, отриманих на його основі.

Третій етап полягав у структурному аналізі ідіолектних засобів на тлі лексичного та морфологічного рівнів мови, структурного роззначення творів. За допомогою методу словникових дефініцій з'ясовано структуру значення лексичних одиниць. Контекстний метод послугував для опису смислів, актуалізованих у текстовій канві художнього твору.

Четвертий етап дослідження полягав у статистичному та квантитативному опрацюванні лексики (для підрахунку частоти вживання різнорівневих мовних одиниць; для встановлення повторюваності, домінування, співвідношення певних мовних одиниць у новелах Василя Стефаника, які є характерними рисами ідіолекту письменника).

На **п'ятому етапі** застосовано порівняльний метод – для визначення спільних і відмінних рис, притаманних мові Василя Стефаника та художній українській прозі загалом чи певним письменникам зокрема. Результати дослідження на останніх етапах слугують матеріалом для укладання статистичного профілю письменника та для здійснення порівняльних досліджень із мовою творів інших письменників.

У **третьому розділі „Основні засади побудови структурно-інформаційної моделі словника мови новел Василя Стефаника”** описано елементи структурно-інформаційної будови словника мови новел із прикладами оформлення реєстрових слів; продемонстровано практичне застосування колокаційного підходу до визначення провідного значення багатозначного слова серед найчастотніших іменників у загальнолітературній та діалектній лексиці новел; здійснено порівняння отриманих фактів із корпусом текстів сучасників новеліста, а також частотного словника сучасної української художньої прози (п. 3.1); статистичні підрахунки, виконані на матеріалі корпусу текстів новел, містять інформацію про кількість слововживань, кількість словоформ (зокрема вжитих один раз і 10 і більше разів), коефіцієнт різноманітності, середню повторюваність, коефіцієнт винятковості та коефіцієнт концентрації, співвідношення частин мови, загальноновживаної / діалектної, авторської / прямої мови (персонажів) у творах та їх порівняння із показниками окремих авторів (сучасників Василя Стефаника та тих, що писали в середині ХХ століття) та дали змогу виокремити риси, притаманні ідіолекту новеліста (п. 3.2).

Структурно-інформаційна модель словника мови новел Василя Стефаника є трирівневою. Мікроструктуру словника представлено лівобічною частиною (де розміщено лему; у випадку омонімів – із відповідною числовою позначкою) та правобічною частиною, де розміщено інформацію про частоту вживання; граматичний / стилістичний коментар (де необхідно); тлумачення; цитування з відповідною паспортизацією; словозмінну парадигму із частотою вживання у кожному

творі (зі вказівкою на твір). На рівні макроструктури до реєстру увійшли всі повнозначні (лематизовані) та службові частини мови.

Для побудови правобічної частини словникової статті, зокрема зони тлумачень, задіяно методи кількісного, статистичного (із застосуванням статистичних показників ступеня зв'язності елементів колокацій *t-критерій* та *взаємної інформації*) і контекстного аналізу, що дає змогу точніше окреслити порядок подання тлумачень багатозначних слів. Наприклад, розглянуто лему „хата”, яка згідно з СУМ-11 має 5 значень („1) *сільський одноповерховий житловий будинок, домівка, господа, тимчасове пристановище для кого-небудь*; 2) *внутрішнє житлове приміщення такого будинку; кімната*; 3) *родина, люди, які живуть, перебувають в одному такому приміщенні*. 4) *розм. квартира (у 1 знач.)*; 5) *Уживається у складі назв деяких сільських установ*” (СУМ-11, XI том, ст. 29). Застосування показників ступеня асоціації колокацій *t-критерій* та *взаємної інформації* довело, що в новелах найстійкіший зв'язок цієї лєми з просторовими прийменниками (*в / у, до*), що вказує на те, що першочерговим тлумаченням у словнику мови новел буде друге значення з СУМ-11 („*внутрішнє житлове приміщення сільського одноповерхового житлового будинку*”). Відтак словникова стаття лєми „хата” матиме такий вигляд:

ХАТА (249) ім., жін. 1) внутрішнє житлове приміщення сільського одноповерхового житлового будинку, домівка, господа: *Як був пеній, то ніколи не входив до хати, ци зима, ци літо – все на присні спав.* (З міста йдучи: 26). *Отсе така хвиля, коли малі діти вибігають із хати на толоку і граються дуже весело, нервово.* (Вечірня Година: 80); *Діти бігали по хаті.* (Осінь: 71); 2) родина, люди, які живуть, перебувають в сільському житловому будинку. – *То я, стара, маю вам хати звеселювати?* (Басараби: 146)

у порівн. (1) *На такі шіфі як хата великі, таким морем шіроким, шіроким, гет, гет ...* (Мамин Синок: 65)

♦ **повна хата (3)** велика кількість, дуже багато кого-, чого-небудь. *Гостей у Івана повна хата, газди і газдині.* (Камінний Хрест: 88)

Встановлено переважання структур простих іменних порівняльних зворотів зі словом *як* (загальною кількістю 68 зворотів), де порівнюють предмети та ознаки. Контекстний аналіз застосовано для виокремлення компонентів порівняльної конструкції *S як O* (того, що порівнюють (компарандум, S), з чим порівнюють (компаратум, O), ознаки, за якою порівнюють) та визначення належності компарандума до певного концепту в мовній картині світу Під час роботи виявлено, що переважають конструкції, де найчастіше порівнюваними є слова на позначення людини як живої істоти (частини тіла / фізичні дані / зовнішність): *руки як зрєбло,*

волосе як лен / молоко. Порівнявши відомості, отримані з корпусу творів сучасників новеліста, лише в новелах Василя Стефаника фіксуємо порівняльні звороти *біле як рантух, лице як шкіра на старім чоботі, міцних і здорових, як ковбки, тямбу нехарапутну, як у старої конини*, які становлять індивідуально-авторські конструкції у творчості Василя Стефаника.

Фразеологізми описано при основних компонентах або за певними формальними ознаками (при першому повнозначному компонентові, за граматично основним компонентом) та розташовано в кінці словникової статті відповідного слова після знака ромба. Тлумачення подають згідно з традиційною практикою: іменні – іменним описовим зворотом; дієслівні – інфінітивним описовим зворотом, ад’єктивні – прикметниковим описовим зворотом; адвербіальні – прислівниковим описовим зворотом. Якщо слово використано в кількох фразеологізмах, то їх подаємо в переліку з абзацу, з нумерацією значень і кількістю вживань у дужках.

Кількість словоформ, які вжито один раз, становить 15,63 % усього тексту, а індекс винятковості – 0,156, який слугує показником варіативності лексики. Протилежним до індексу винятковості є індекс концентрації – відношення кількості слів із частотою 10 і вище (у Василя Стефаника це 647 слів) до обсягу тексту (у Василя Стефаника – 53978 слововживань), який становить 0,01. Відносно невелика кількість високочастотної лексики (і, відповідно, низький індекс концентрації) засвідчує різноманіття лексики новел Василя Стефаника.

Окремо досліджено частоту вживання слів за частинами мови в кожній новелі, а також сумарно в усіх новелах. На думку С. Бук, Н. Дарчук, А. Середницької, такі дослідження в когнітивно-лінгвістичному аспекті перспективні для побудови мовної картини світу письменника та ідеографічного опису мови його творів. Отримані відомості порівняно з показниками, властивими сучасній українській прозі, що оприявило доволі рівномірний розподіл: іменники – 23,34 % у новеліста та 25,8 % у сучасній українській художній прозі; дієслова – 20,6 % та 18,3 %; займенники – 12,79 % та 11,0 %; прислівники – 6,08 % та 8,4 %; числівники – 0,99 % та 1,0 %; службові частини мови – 31,46 % та 26,8 % відповідно, окрім прикметників – 4,88 % та 8,6 % відповідно.

Розподіл слів за частинами мови у Василя Стефаника вказує на переважну кількість дієслів (1594 словоформ), далі йдуть іменники (777 словоформ) і прикметники (688 словоформ). Ті самі відомості лягли в основу аналізу показників індексу епітетизації, дієслівних означень і номінальності. Виявлено, що в новелах Василя Стефаника індекс епітетизації вищий (2,34), а найнижчим виявився показник дієслівних означень (0,30).

Порівняльний аналіз лексики Василя Стефаника з творами письменників того ж хронологічного періоду, а також показниками сучасної української художньої прози, виявив, що вона містить відмінності й у розрізі загальномовного слововживання (вживання деяких лексем відсутні в сучасній українській художній прозі (наприклад, слово *жид*) або значно менше кількісно (наприклад, *дитина*, *баба*, *Бог*, *жінка*, *людина*, *мама*, *піч*, *мужик*, *церква*, *гріх* та *тато*).

Вивчення частки діалектної лексики довів, що вона становить близько 22 % усієї лексики та має рівномірний розподіл по новелах та в хронологічно різні періоди творчості, що засвідчує те, що впродовж творчої діяльності Василя Стефаника не відступав від своєї творчої манери та послуговувався однаковим стилем. Тематичний розподіл діалектизмів показав, що переважальною є лексика на позначення людини (211 лексичних одиниць), що засвідчує антропоцентричність творчості новеліста. Наступними йдуть слова на позначення різних аспектів тогочасного побуту сільського населення Покуття (страви та напої, предмети побуту, а також назви будівель та конструкцій), що становить разом 258 лексичних одиниць.

Частка діалектної лексики в авторській мові – 2 %, а в прямій – 10 %. Отже, мова персонажів у новелах Василя Стефаника містить більшу частку діалектної лексики, що доводить використання в прямій мові персонажів саме такого мовлення, яке було притаманне мешканцям описаного регіону.

Додатково визначено відсоток авторської та прямої мови в текстах новел (37% і 63 % відповідно), що засвідчує високий ступінь діалогічності творів письменника.

Показник ентропії символічного тексту як „кількості інформації” (за А. Колмогоровим та К. Шенноном), яку несуть символи без урахування змісту тексту, ефективний для визначення авторства твору в разі невеликого обсягу досліджуваного тексту або ж за відсутності традиційних об’єктів дослідження мови автора. Застосування методу аналізу частотності біграм і їхньої ентропії як формальних методів зіставлення особливостей стилю письменників одного хронологічного періоду (П. Карманського, В. Стефаника, М. Яцкова), але різних жанрів (новел і поезії) дало підстави для висновку, що ентропія біграм у всіх масивах текстів має приблизно однакове значення (із незначною відмінністю у Василя Стефаника), незважаючи на зіставлення поетичних творів із прозовими. Установлено, що вирахунок ентропії біграм не може бути надійним методом визначення авторського інваріанта, а для статистичних досліджень краще робити вибірки у випадковий, а не послідовний спосіб.

Квантитативний аналіз символів (графем) у новелах Василя Стефаника для порівняння милозвучності новел із відповідним

показником художнього стилю підтвердив його значення найближчим до художньої прози української літератури (0,79 та 0,793 відповідно).

Порівняння даних про довжину слів і речень у новелах Василя Стефаника (57 новел) з аналогічними показниками у творчості інших українських письменників, зокрема Михайла Яцкова (42 новели), Романа Іваничука (43 новели), які представляють один жанр, але хронологічно різні періоди, показало, що: 1) середня довжина слів у новелах Василя Стефаника (4,29) найближча до творів Андрія Головка (4,74), Олександра Довженка (4,73) та Михайла Яцкова (4,75); 2) середня довжина речень досліджуваних авторів відповідає показникам художньої прози української мови. Загалом розподіл слів у реченнях новел Василя Стефаника та Михайла Яцкова збігається – речення в діапазоні 4–12 слів становлять більше половини обсягу текстів новел (58,4 % та 56,9 % відповідно проти 42,6 % у Романа Іваничука), що можна пояснити хронологічно близьким періодом творчості Василя Стефаника та Михайла Яцкова, а також підтипом жанру модерністської новели, властивим їм обом. Тяжіння новелістів до коротких речень виразає лаконічність їхнього стилю.

ВИСНОВКИ

Дисертація присвячена розпрацюванню структурно-інформаційної моделі словника мови новел Василя Стефаника, майстра експресіоністичної новели, базованого на корпусі їх прижиттєвих видань.

Теоретико-методологічним підґрунтям міждисциплінарного підходу до укладання словників мови письменників є напрацювання з авторської лексикографії, корпусної лінгвістики, лінгвістики тексту, досліджень ідіолекту письменника, концептуальної та мовної картин світу особистості, діалектної лексики в художніх творах, статистичного опрацювання корпусів текстів. Застосування такого підходу уможливило поєднання відомостей та структури конкордансу, тлумачного, і частотного словників, а також інформації, яку зазвичай подають у глосаріях та словниках фразеологізмів письменника, у єдиній лексикографічній праці.

Створений дослідницький лінгвістичний корпус новел Василя Стефаника є референційним; синхронним; статичним; спеціальним; повнотекстовим; писемним; одномовним та частково роззначеним.

Здійснений аналіз символів (графем) указав, що милозвучність новел відповідає значенню цього показника найближчим до показника художньої прози сучасної української літератури. Вивчення частотності біграм (двох символів) і їхньої ентропії довів, що цей показник у досліджуваних масивах текстів має приблизно однакове значення, незважаючи на наявність поетичних творів паралельно із прозовими.

Також встановлено, що визначення показника ентропії біграм не є надійним методом у виокремленні авторського інваріанта.

Виокремлення слів і словоформ уможливило отримання інших статистичних характеристик тексту (кількість слів і словоформ, що вжиті один раз, кількість слів і словоформ, що вжиті 10 і більше разів, коефіцієнт різноманітності, коефіцієнт винятковості, коефіцієнт концентрації), які засвідчують різноманіття лексики новел.

Порівняння кількісних відомостей за частинами мови із показниками української художньої прози показало доволі рівномірний розподіл. Серед показників індексів епітетизації, дієслівних означень та номінальності у новелах Василя Стефаника індекс епітетизації вищий, а найнижчим виявився показник дієслівних означень, що корелює із даними сучасної української художньої прози загалом.

Порівняльний аналіз лексики новел виявив, що вона містить відмінності не лише у розрізі загальномовного слововживання (вживання деяких лексем відсутнє в сучасній художній прозі або значно менше кількісно), а й у переважанні певних тематичних груп лексики (переважальна кількість діалектних лексем на позначення людини та її оточення / побуту).

Діалектний словник засвідчив рівномірний розподіл по новелах і в різні періоди творчості письменника. Відсоток авторської та прямої мови у текстах новел указує на високий ступінь діалогічності новел.

Статистичні обчислення довжин слів і речень у новелах та порівняння цих відомостей із аналогічними показниками інших українських авторів засвідчили, що за довжиною слів новели Василя Стефаника є найближчими до творів Андрія Головка, Олександра Довженка та Михайла Яцкова, а розподіл довжини речень збігається в новелах Василя Стефаника та Михайла Яцкова, що можна пояснити близьким у часі періодом творчості обох письменників, а також піджанром новели, характерним для обох авторів. Лаконічність стилю Василя Стефаника підтверджено тяжінням до коротких речень.

Переважальною у корпусі новел є модель простих іменних порівняльних конструкції типу *S як O*. Контекстний аналіз показав, що найчастіше порівнюваними (*S*) є слова на позначення людини як живої істоти.

Перспективним вважаємо ґрунтовний аналіз лексики новел Василя Стефаника на граматичному та семантичному рівнях, а також розширення джерельної бази корпусами текстів епістолярної спадщини новеліста, паралельними корпусами різних видань і перекладів новел. Отримані відомості сприятимуть аналізу лексики у порівнянні з іншими письменниками, зокрема сучасниками новеліста, та інтегральному опису

лексики творів Василя Стефаника і побудові ідеографічного словника мови новел у перспективі.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці у наукових фахових виданнях України

1. Данчевська Ю. О., Кульчицький І. М., Ліхнякевич І. О. Деякі аспекти використання лінгвістичних корпусів слов'янських мов. *Науковий вісник Волинського національного університету імені Лесі Українки. Серія „Філологічні науки. Мовознавство”*. Луцьк, 2012. С. 216–220.
2. Данчевська Ю. О., Кульчицький І. М., Ліхнякевич І. О. Деякі аспекти створення та використання паралельних корпусів. *Науковий вісник Волинського національного університету імені Лесі Українки. Серія „Філологічні науки. Мовознавство”*. Луцьк, 2013. С. 48–52.
3. Данчевська Ю. О., Кульчицький І. М., Ліхнякевич І. О. Лексикографічні аспекти опису мови письменників. *Наукові записки Кіровоградського державного педагогічного університету. Серія „Філологічні науки”*. Кіровоград, 2013. С. 154–158.
4. Калимон Ю. О., Кульчицький І. М., Ліхнякевич І. О. Ідіолект, ідіостиль, індивідуальний стиль. Тотожне чи різне? *Науковий вісник Східноєвропейського національного університету імені Лесі Українки. Серія „Філологічні науки”*. Луцьк, 2014. С. 226–229.
5. Калимон Ю. О., Кульчицький І. М. Кількісний аналіз новел В. Стефаника та їх перекладів англійською мовою. *Науковий вісник Східноєвропейського національного університету імені Лесі Українки. Серія „Філологічні науки”*. Луцьк, 2015. С. 193–197.
6. Калимон Ю. О. Використання квантитативних методів у дослідженні ідіолекту письменника. *Науковий вісник Східноєвропейського національного університету імені Лесі Українки. Серія „Філологічні науки”*. Мовознавство. Луцьк, 2017. С. 417–421.
7. Калимон Ю. О. Співвідношення термінів на позначення системи мовних засобів письменника. *Актуальні проблеми філології та перекладознавства*. Хмельницький, 2017. Вип. 12. С. 92–96.
8. Калимон Ю. О., Кульчицький І. М., Ліхнякевич І. О. Ентропія біграм як метод формального аналізу художніх творів. *Актуальні питання іноземної філології*. Луцьк, 2018. С. 105–111.
9. Калимон Ю. О. Комп'ютерна лексикографія: виклики та перспективи. *Актуальні питання іноземної філології*. Луцьк, 2019. С. 112–118.

10. Kalymon Yu. O. Narrative and Language Types of Short Stories by Vasyl Stefanyk: Quantitative Approach. Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Філологія. Соціальні комунікації. Том 31 (70) № 2. Київ, 2020. С. 125-128.

Наукова праця у періодичному виданні іншої держави

11. Калимон Ю. О. Квантитативне дослідження графем новел В. Стефаника. *Матеріали ІХ Оломоуцького симпозиуму українців Середньої та Східної Європи*. Оломоуць, 2018. С. 66–70. ISSN 0231-634X

Наукові праці апробаційного характеру

12. Данчевська Ю. О., Кульчицький І. М. Деякі аспекти створення корпусу художніх творів В. С. Стефаника. *Матеріали Міжнародної наукової конференції MegaLing 2012 „Прикладна лінгвістика та лінгвістичні технології”*. Київ, 2013. С. 143–149.

13. Данчевська Ю. О., Кульчицький І. М., Ліхнякевич І. О. Lexical and Statistical Aspects of V. Stefanyk's Idiolect Study. *Матеріали VIII Міжнародної наукової конференції „Комп'ютерні науки та інформаційні технології” / CSIT 2013*. Львів : Видавництво Львівської політехніки, 2013. С. 34–35.

14. Калимон Ю. О. Розробка електронного словника: ключові етапи та виклики. *Збірник наукових праць ІV Міжнародної наукової конференції „Людина. Комп'ютер. Комунікація” / за ред. О. П. Левченко*. Львів : Видавництво Львівської політехніки, 2019. С. 80–83.

Праці, які додатково відображають наукові результати дисертації

15. Данчевська Ю. О. Мова творів В. С. Стефаника: сучасний стан та перспективи дослідження. *Людина. Комп'ютер. Комунікація: Збірник наукових праць / за ред. О. П. Левченко*. Львів : Видавництво Львівської політехніки, 2013. – С. 44–46.

16. Калимон Ю. О. Абзац як категорія структурного поділу тексту. *Збірник наукових праць „Лінгвокомп'ютерні дослідження”*. Вип. 11. Вінниця, 2018. С. 135–140.

17. Калимон Ю. О. Моделювання: лінгвістичний аспект. *Науковий журнал „Молодий вчений”*. ДДПУ ім. Івана Франка. Дрогобич, 2019. С. 88–91.

АНОТАЦІЯ

Калимон Ю. О. Структурно-інформаційна модель словника мови новел Василя Стефаника. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата філологічних наук за спеціальністю 10.02.01 – українська мова. – Національний університет „Львівська політехніка”, Чернівецький національний університет імені Юрія Федьковича, Міністерство освіти і науки України, Чернівці, 2020.

У дисертації розроблено структурно-інформаційну модель словника мови новел Василя Стефаника на основі корпусу текстів останніх прижиттєвих видань його творів; поглиблено вивчення західноукраїнського варіанта української літературної мови, який існував на зламі XIX та XX століть; здійснено порівняльне дослідження різнорівневих мовних одиниць із показниками української прози загалом і мови деяких письменників зокрема. У праці доповнено і розвинуто наукові уявлення про вивчення ідіолекту письменника з застосуванням корпуснобазованих технологій і статистичного апарату. Розроблено структурно-інформаційну типологію словників, проаналізовано й описано інформаційне наповнення наявних українських словників мови письменників різних типів. Охарактеризовано різноманіття підходів до трактування термінів на позначення мовно-виражальних засобів письменників, методів їх корпуснобазованого та статистичного опрацювання.

У праці вперше досліджено мовну структуру новел Василя Стефаника на рівні звуків, біграм, розподілу довжини слів, розподілу кількості слів у реченнях. Запропоновано застосування статистичних показників ступеня асоціації слів для визначення провідного значення полісемантичного слова.

Ключові слова: авторська лексикографія, структурно-інформаційна модель словника, конкорданс, ідіолект, частотний словник, корпус текстів, статистична лінгвістика.

АННОТАЦІЯ

Калымон Ю. А. Структурно-информационная модель словаря языка новелл Василия Стефаника. – На правах рукописи.

Диссертация на соискание учёной степени кандидата филологических наук по специальности 10.02.01 – украинский язык. – Национальный университет „Львовская политехника”, Черновицкий национальный университет имени Юрия Федьковича, Министерство образования и науки Украины, Черновцы, 2020.

В диссертации разработана структурно-информационная модель словаря языка новелл Василия Стефаника на базе корпуса текстов последних прижизненных изданий его произведений; изучен западноукраинский вариант литературного языка рубежа XIX и XX веков.

В работе дополнено и развито научное представление об изучении идиолекта писателя на основе исследовательского лингвистического корпуса и с применением статистического анализа; проведено сравнительное исследование разноуровневых языковых единиц в новеллах Василия Стефаника с показателями украинской художественной прозы в целом и некоторых писателей в частности. Разработана структурно-информационная типология словарей, проанализировано и описано информационное наполнение существующих украинских словарей языка писателей разных типов. Исследовано многообразие подходов к трактовке терминов для обозначения выразительных средств языка писателей, методов их корпусной и статистической обработки.

В работе впервые исследована языковая структура новелл Василия Стефаника на уровне звуков, биграмм, распределения длины слов и количества слов в предложениях. Предложено применение статистических показателей степени связности слов для определения ведущего значения полисемантического слова.

Ключевые слова: авторская лексикография, структурно-информационная модель словаря, конкорданс, идиолект, частотный словарь, корпус текстов, статистическая лингвистика.

ABSTRACT

Kalymon Yu. O. Structural and informational dictionary model of Vasyl Stefanyk's short stories language. – Manuscript.

Thesis for a Candidate Degree in Philology (PhD). Speciality 10.02.01 – Ukrainian language. – Lviv Polytechnic National University, Yuriy Fedkovych Chernivtsi National University, Ministry of Education and Science of Ukraine, Chernivtsi, 2020.

This thesis focuses on the development of an informational and structural model of the vocabulary of Vasyl Stefanyk's language based on the text corpus of the latest lifetime editions of his short stories, thus deepening the study of the Western Ukrainian version of the Ukrainian literary language, which existed at the turn of the 19th and 20th centuries. The study has supplemented and developed scientific ideas about the study of the idiolect of the writer by means of corpus-based technologies and statistical apparatus. A comparative study of linguistic units at different levels with indicators of the Ukrainian prose in general and some writers in particular.

The research corresponds to a modern tendency aimed at a comprehensive study of language units' functions at different levels in the literary text, constant growth of scientific interest in integration of different methods and approaches to their study. A structural and informational typology of dictionaries has been developed, the information content of existing Ukrainian dictionaries of the

language of writers of different types has been analyzed and described. A variety of approaches to the interpretation of terms for the designation of expressive means of the language of writers, methods of their corpus and statistical processing are investigated.

For the first time, distribution of words according to parts of speech, literary and dialectal lexis, author's and character distribution of vocabulary was quantitatively analyzed; the indices of epithetization, verb designation and nominality are calculated. The thematic distribution of dialectal vocabulary of nouns is offered and their leading concepts regarding the linguistic picture of the world (according to V. Zhayvoronok) are outlined. For the first time, the coefficient of exclusivity of the vocabulary of short stories is determined; a comparative analysis of the use of color names and comparative noun constructions with the word *як* ("yak / as, like") in short stories and works of the writer's contemporaries is performed. Collocational approach by means of association measures (which determine how strong the connection between words is), namely T-score and Mutual Information, are to reveal the leading meaning of a polysemous word.

Firstly conducted statistical studies of different levels of language and text (symbols, sounds, bigrams, words and word forms length and number, sentences length) indicated the closeness of short stories published during the life of Vasyl Stefanyk to the indicators of Ukrainian modern fiction in general and some individual writers. We consider a thorough analysis of the vocabulary of Vasyl Stefanyk's short stories at the grammatical and semantic levels, as well as the expansion of the source base with corpora of texts of the novelist's epistolary, parallel corpora of different editions and translations of short stories, as perspectives of further studies. The obtained information promotes the analysis of vocabulary in comparison with other writers, contemporaries of the novelist in particular, and integral description of vocabulary of Vasyl Stefanyk's works and compilation of an ideographic dictionary of the of short stories language units in perspective.

Key words: author's lexicography, structural and informational model of a dictionary, concordance, idiolect, frequency dictionary, text corpus, statistical linguistics.

Підписано до друку 02.11.2020 р.
Формат 60×90 1/16. Папір офсетний.
Друк на різнографі. Умовн. друк. арк. 1,5. Обл.-видав.арк. 0,89.
Тираж 100 прим. Зам. 200944.

Поліграфічний центр
Видавництва Національного університету “Львівська політехніка”
вул. Ф.Колесси, 4, 79013, Львів
Реєстраційне свідоцтво серії ДК № 4459 від 27.12.2012 р.