

УДК 811.22'374'373'373.72-26 : 821.22-1.02'19". Ф. Фаррохзад

ЧАСТОТНИЙ СЛОВНИК МОВИ ТА ЛЕКСИКОСТАТИСТИЧНІ ОСОБЛИВОСТІ ІДІОСТИЛЮ ФОРУГ ФАРРОХЗАД

Ірина Яремчук

*Львівський національний університет імені Івана Франка,
кафедра сходознавства,
вул. Університетська, 1/239, Львів, Україна, 79602,
тел.: (032) 239 47 04, e-mail: irynayaremchuk@gmail.com*

Проаналізовано ідіостиль іранської поетеси та кінорежисера ХХ століття, представниці літературної течії “Ше’ре ноу” Форуґ Фаррохзад у статистичному аспекті. Подано особливості укладання частотного словника мови Форуґ Фаррохзад та списку 200 найчастотніших слів її поезій.

Ключові слова: ідіостиль, частотний словник мови поетеси, лематизація слів, статистичні характеристики словника мови письменника.

У сучасній лінгвістиці усе частіше залучають кількісні методи для аналізу мовних одиниць. Статистичний аналіз тексту у вивченні ідіостилю письменника застосовував М. Мухін для дослідження творчості російських письменників (М. Булгакова, В. Набокова, А. Платонова та М. Шолохова) [5–8]. Активно створюються словники мови письменників, метою яких є відображення мовних особливостей, притаманних ідіостилю письменника чи поета. В іраністиці нам відомі частотні словники мови Унсурі [9], Гафеза [16], частотний словник мови “Шахнаме” [13]. Застосування статистичного методу для визначення набору ключових слів поезій Форуґ Фаррохзад дає змогу охарактеризувати кількісні параметри її ідіостилю, оскільки кількісна характеристика мовних фактів разом із якісною допомагає різносторонньо описувати мовні явища.

Для опису концептуального простору поезій Форуґ Фаррохзад (1935–1967), відомої іранської поетеси та кінорежисера, представниці літературної течії “Ше’ре ноу”, та його структурування ми використали, окрім концептуального, статистичний метод. Укладання частотного словника мови та списку 200 найчастотніших слів поезій Форуґ Фаррохзад дало змогу на основі переліку концептів, лексичні репрезентанти яких увійшли до цього списку, охарактеризувати концептуальний простір поезій, виділити у ньому концептуальні поля СВІТ ЛЮДИНИ та ФІЗИЧНИЙ СВІТ, які ми поділили на підкласи та кластери [про це див. : 11]. Частотний словник мови Форуґ Фаррохзад та застосування статистичного методу допомогли охарактеризувати лексикостатистичні особливості ідіостилю поетеси, що стало метою цієї статті.

Усі поетичні твори іранської поетеси – чотири збірки *اسیر* Asir “Полонянка” (1955), *دیوار* Divār “Стіна” (1956), *عصیان* ‘Asyān “Бунт” (1958), *تولدی دیگر* Tavvalodi digar “Нове народження” (1963) та сім поем збірки *ایمان بیاوریم به آغاز فصل سرد* Imān beyāvarim be

āqhāz-e fasl-e sard... “Повіримо у настання холодів...” [17] стали генеральною сукупністю обсягом 3171 лексична одиниця. За допомогою статистичного методу було підраховано абсолютну частоту слів, вжитих у віршах поетеси. Ми уклали частотний словник мови Форуг Фаррохзад, у якому слова розміщені за алфавітом (відповідно до правописного словника перської мови [15]), біля кожного слова було вказано абсолютну частоту вживання у творах поетеси [12, с. 176–248]. У процесі укладання частотного словника мови Форуг Фаррохзад було використано принципи укладання частотного словника газетної лексики сучасної перської мови [3; 4]. Варто зауважити, що різні варіанти одного і того ж слова (фонетичні, розмовні), ламану арабську множину підраховано разом, а варіанти подано у дужках (наприклад, راه و خاموش; راه و خاموش). Окремо не фіксувалися граматичні форми і повнозначні морфеми – основи теперішнього часу простих і прийменникових дієслів та дієслівних напівафіксів, оскільки фіксувалися лексеми, а не похідні від них словотвірні чи граматичні елементи. Окремо підраховано субстантивовані дієслівні форми, зафіксовані у словниках перської мови [14]. Складні слова вважалися неподільною одиницею. Власні назви фіксувалися як звичайні слова. Похідні іменники із суфіксом -ی (یای مصدری) та прикметники з суфіксом -ی (یای نسبت) фіксувалися як окремі лексеми. Омоніми розмежовано арабськими цифрами, відповідно до принципів подання омонімів у перських тлумачних словниках [14].

Лематизацію, або виведення заголовної форми слова, виконували відповідно до принципів лематизації частотного словника газетної лексики сучасної перської мови [3], із врахуванням того факту, що фіксувалася лексика ідіостилю автора. При фіксації дієслів заголовною формою стала стверджувальна форма інфінітива; зводилися часові, особові, заперечні форми та форми наказового способу; складні дієслова вважали складними словами і реєстрували як окреме слово; префіксальні дієслова реєстрували окремо як одне слово; дієприкметники, каузативні форми фіксували окремо; якщо дієслово вживається й у правильній, і у неправильній формі, то його частотність сумували, а менш вживаний варіант подано у дужках; зв'язка است ast 'e' вважалася формою дієслова بودن budan 'бути'; випадки її вживання як формотворчого афікса не фіксували. Для іменників заголовною формою стала форма однини, у визначеному стані; зводилися форми регулярної і арабської ламаної множини. У дужках біля відповідних іменників подано зафіксовані форми арабської множини; частоту множинних іменників підраховано окремо; слова, які вживаються і як іменники, і як прикметники чи інші частини мови розглянуто як полісемантичні лексеми й усі їхні значення зареєстровано під одним заголовним словом. Під час реєстрації прикметників та прислівників зводилися вищий і найвищий ступені порівняння; слова, утворені за моделлю “префікс + іменник” чи “префікс + بی + іменник” і інші зафіксовано як окремі лексеми. Для займенника реєструвалися усі самостійні форми – особові, вказівні, означальні та інші займенники. Займенникові енклітики, які не мають самостійності окремого слова, не реєструвалися. При фіксації числівника складені та дробові числівники зафіксовано покомпонентно окремо, оскільки вони утворюються за правилами; частоту вживання порядкових числівників підраховано окремо; числівники, які записані цифрами, із тексту вилучали. Відіменникові прийменники зафіксовано як окремі лексеми. Усі інші частини мови (післяйменник, вигук, сполучник, частка) реєстрували як окремі лексеми.

Наступним кроком стало укладання частотного списку слів поезій, у якому слова були розташовані за спадом частот, подано абсолютну частоту слова в усій творчості поетеси та дані окремо за збірками, що допомогло аналізувати мовні особливості поезії Форуг Фаррохзад також і в еволюційному аспекті. До списку 200 найчастотніших слів поезій Форуг Фаррохзад увійшли лише повнозначні слова, оскільки найбільшу частотність мали службові слова і займенники, а частота їх вживання є малопродуктивною для аналізу семантики тексту, тому вони були вилучені зі списку (до уваги також був взятий закон Гіро – найчастотніші слова (а до першої десятки списку, як уже згадувалося вище, увійшли службові слова та займенники) та становлять значну частину тексту, проте їх інформаційна вартість є невисокою, інформативними є рідковживані слова [10, с. 114]). Для наочності подаємо частину списку у таблиці 1.

Таблиця 1

Список 200 найчастотніших слів у поезіях Форуг Фаррохзад

Частота вживання у збірці					Абсолютна частота	Слово	№
“Повіriamo у настання холодів...”	“Нове народження”	“Бунт”	“Стіна”	“Полонянка”			
1	2	3	4	5	6	7	8
8	18	16	28	91	161	دل	1
16	55	11	26	41	149	شب	2
6	30	18	30	56	140	چشم	3
18	22	11	20	56	127	گفتن	4
12	24	10	19	54	119	عشق	5
11	45	13	3	43	115	رفتن	6
8	35	15	21	32	111	دیدن	7
1	17	25	31	25	99	راه (راه	8
28	27	9	7	26	97	آمدن	9
-	7	9	25	53	94	لب	10
9	31	15	11	25	91	دست	11
12	38	16	6	14	86	روز	12
29	28	16	2	6	81	صدا	13
4	13	14	15	34	80	نگاه (نگاه	14
-	9	18	18	23	77	خاموش (خاموش، خاموش	15
8	22	10	11	18	69	خواب	16
-	14	16	15	22	67	سینه	17
4	15	8	11	26	64	پا	18
11	25	3	10	13	62	گل	19
9	35	3	3	11	61	آب	20

Застосування статистичного методу допомогло зробити висновки про різноманітність словника поетичної мови Форуґ Фаррохзад, його багатство.

У всіх поетичних творах Форуґ Фаррохзад, про що уже йшлося вище, зафіксовано 3171 слова, що склало обсяг словника поетеси (V).

Словник поетеси є доволі різноманітним, оскільки уже до першої тисячі увійшли слова, які мали абсолютну частоту 3, а починаючи із 1797-го слова абсолютна частота становила 1. Багатство словника поетичної мови Форуґ Фаррохзад, індекс різноманітності розраховано за такою формулою

$$B = V / N,$$

де B – багатство словника, V – обсяг словника, N – обсяг тексту [1, с. 30].

Отримано такі результати – 0,072 (для порівняння – у частотному словнику української публіцистики – 0,070 [1, с. 30]).

Також важливим для характеристики поетичної мови Форуґ Фаррохзад є індекс винятковості, що характеризує варіативність лексики, тобто частку словника, яку займають слова, що трапились у тексті 1 раз. Використовуючи формулу

$$E_c = V_1 / V,$$

де E_c – індекс винятковості для словника, V_1 – кількість лексем із частотою 1, а V – загальна кількість лексем [1, с. 30], ми отримали індекс винятковості 0,433 (для порівняння – індекс винятковості частотного словника офіційно-ділового стилю української мови становить 0,280 [2, с. 166–172]).

Для характеристики концентрації словника Форуґ Фаррохзад вираховано індекс концентрації словника – відношення кількості слів у словнику з абсолютною частотою 10 і більше до загальної кількості слів у словнику:

$$V_{10} / V,$$

де V_{10} – кількість слів у словнику з абсолютною частотою 10 і більше, а V – загальна кількість слів у словнику [1, с. 36].

Отримано дані 0,146 (для порівняння – частотний словник поетичної мови українських поетів – 0,098 [1, с.36]).

Укладений частотний словник мови Форуґ Фаррохзад допоможе характеризувати мовні особливості ідіостилю поетеси. Застосування статистичного методу дало змогу охарактеризувати кількісно мовні факти ідіостилю Форуґ Фаррохзад, а саме – різноманітність словника її поетичної мови, варіативність лексики, багатство словника поетичної мови.

Подальші перспективи досліджень полягають у можливості застосування концептуального аналізу (поєднання частотного, контекстологічного і концептуального напрямку вивчення тексту) до поетичних творів іранських поетів різних періодів, зіставного дослідження ідіостилів поетів Ірану.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Бук С. Н. Основи статистичної лінгвістики: навчально-методичний посібник / С. Н. Бук / [відп. ред. проф. Ф. С. Бацевич]. – Львів : Видавничий центр ЛНУ імені Івана Франка, 2008. – 124 с.

2. Бук С. Статистичні характеристики лексики основних функціональних стилів української мови: спроба порівняння / С. Бук // Лексикографічний бюлетень. – 2006. – Вип. 13. – С. 166–172.
3. Максимів О. Принципи укладання частотного словника газетної лексики сучасної перської мови / О. Максимів // Вісник Київ. ун-ту імені Т. Шевченка.
4. Максимів О. Й. Методика виявлення ключових слів на основі зіставлення лексики частотних словників / О. Й. Максимів // Studia Linguistica : зб. наук. праць. – К., 2009. – Вип. 3. – С. 192–199.
5. Мухин М. Ю. Идиостиль автора : квантитативный анализ и его филологическая интерпретация [Электронный ресурс] / М. Ю. Мухин // Материалы “Диалога-2005”. – Режим доступа : <http://www.dialog-21.ru/Archive/2005/MukhinM/MukhinM.htm>.
6. Мухин М. Ю. Лексические универсалии художественного текста : статистика и идеография / М. Ю. Мухин // Изв. Урал. гос. ун-та. – Екатеринбург, 2007. – № 53. – Вып. 14. – С. 105–109. – (Сер. 2. Гум. науки).
7. Мухин М. Ю. Проблема автора и статистический анализ текста / М. Ю. Мухин // Литература Урала : история и современность : сб. статей. – Вып. 3. – Т. 1. – С. 132–143.
8. Мухин М. Ю. Общее и индивидуальное в словаре языка писателя / М. Ю. Мухин // Креативная языковая личность в этносоциокультурном прагмалингвистическом контексте : сб. науч. трудов. – Екатеринбург : Изд-во Урал. пед. ун-та, 2008. – С. 143–147.
9. Османов М.-Н. О. Частотный словарь Унсури / М.-Н. О. Османов. – М. : Наука, 1970.
10. Тищенко К. М. Основи мовознавства : системний підручник / К. М. Тищенко. – К. : Видавничо-поліграфічний центр “Київський університет”, 2007. – 308 с.
11. Яремчук І. В. Концептуальний простір поезії Форуг Фаррохзад : автореф. дис. на здобуття наук. ступеня канд. філол. наук : спец. 10.02.13 “Мови народів Азії, Африки, аборигенних народів Америки та Австралії” / І. В. Яремчук; Національна академія наук України. – К., 2010. – 16 с.
12. Яремчук І. В. Концептуальний простір поезії Форуг Фаррохзад : дис. на здобуття наук. ступеня канд. філол. наук : спец. 10.02.13 “Мови народів Азії, Африки, аборигенних народів Америки та Австралії” / І. В. Яремчук; Національна академія наук України. – К., 2010. – 255 с.
13. http://fa.wiktionary.org/wiki/پيوست:فهرست_بسامدی_شاهنامه_فردوسی [Електронний ресурс].
 14. انوری، حسن. فرهنگ روز سخن / حسن انوری. – تهران، سخن، ۱۳۸۳. – ۱۳۷۶ ص.
 15. صادقی، علی اشرف. فرهنگ املائی خط فارسی / علی اشرف صادقی. – تهران، ۱۳۸۶. – ۲۱۹ ص.
 16. صدیقیان م. د. فرهنگ واژه نمای حافظ به انضمام فرهنگ بسامدی / دکتر مهین دخت صدیقیان. – تهران : روزانه، ۱۳۷۸. – ۱۳۵۹ ص.
 17. مجمعه اشعار: فروغ فرخزاد. – تهران: شادان، ۱۳۸۳. – ۳۶۸ ص.

Стаття: надійшла до редакції 15.10.2010
прийнята до друку 15.11.2010

LANGUAGE FREQUENCY DICTIONARY AND LEXICOSTATISTICAL PECULIARITIES OF IDIOSTYLE OF FORUG FARROKHZAD

Iryna Yaremchuk

*The Ivan Franko Lviv National University,
Oriental studies department,
1, Universytets'ka Str., 239, Lviv, Ukraine, 79602
tel.: (032) 239 47 04, e-mail: irynayaremchuk@gmail.com*

Idiostyle of the Iran poetess and film director of XX century, the representative of literary field “She’re nou” Forug Farrokhzad is analyzed in statistical aspect. Peculiarities of compilation of the language frequency dictionary of Forug Farrokhzad and the list of 200 of the most frequently used words in her poetry are described.

Key-words: idiostyle, language frequency dictionary of poetess, lemmatization of words, statistical characteristics of the language dictionary of poetess.

ЧАСТОТНЫЙ СЛОВАРЬ ЯЗЫКА И ЛЕКСИКОСТАТИСТИЧЕСКИЕ ОСОБЕННОСТИ ИДИОСТИЛЯ ФОРУГ ФАРРОХЗАД

Ирина Яремчук

*Львовский национальный университет имени Ивана Франко,
кафедра востоковедения,
ул. Университетская, 1/239, Львов, Украина, 79602,
тел.: (032) 239 47 04, e-mail: irynayaremchuk@gmail.com*

Проанализирован идиостиль иранской поэтессы и кинорежиссера XX века, представительницы литературного направления “Ше’ре ноу” Форуг Фаррохзад в статистическом аспекте. Указаны особенности формирования частотного словаря языка Форуг Фаррохзад и списка 200 самых частотных слов её поэтических произведений.

Ключевые слова: идиостиль, частотный словарь языка поэтессы, лематизация слов, статистические характеристики словаря языка писателя.