

# ГЕНДЕРНІ ОСОБЛИВОСТІ УКРАЇНСЬКОЇ ХУДОЖНЬОЇ ПРОЗИ (ЗА ДАНИМИ АНАЛІЗУ ЧАСТОТНОСТІ КОЛОКАЦІЙ У ТЕКСТАХ 60-80-Х ТА 2000-Х РОКІВ)

Олена Левченко, Ян Бутельський  
Національний університет «Львівська політехніка»

Сьогодні з'явилася низка праць, присвячених дослідженню мовної реалізації фемінності та маскулінності. Зокрема, дослідники здійснюють спроби виявити: **маркери гендеру на різних рівнях мовної системи** [Голубева 2009; Васильєва 2013; Елемешина 2010; Архангельська 2005; Тараненко 2005] та в текстах різної стильової належності [Викторова 2011]; морфологічні особливості мовлення чоловіків і жінок [Комов 2012]; **розглядають динаміку конструювання гендеру в англійськомовному лексикографічному дискурсі** [Гриценко, Сергеева, Барінова 2012]; **аналізують назви осіб жіночої статі в лексикографічному описі та узусі** [Пузиренко 2005]; **лексичну сполучуваність як засіб репрезентації гендеру** (на матеріалі німецької мови) [Шаманська 2012]; **засоби вираження експресивності в жіночих та чоловічих текстах** (на матеріалі узбецької мови) [Пирматова 2013]; оцінку семантику у вторинних номінаціях [Нгуен Суан Тхе 2011]; **англійську метафоричну картину світу** [Тамерьян 2014] та ін. Увагу також приділено **фразеологічним системам**: А. Емірова аналізує відбиття гендеру у фразеології [Эмирова 2002]; М. Єрмолаєва досліджує фразеологічні неологізми в мовленні чоловіків і жінок, їхні творчі мовленнєві здібності на матеріалі російської мови [Ермолаева 2008]; результати дослідження фразеологізмів із маскулінною гендерною референцією в англійській та російській мовах представлено у праці [Арсентьева, Коноплёва, 2009]; гендерну маркованість емотивних стійких сполук в осетинській мові [Гацалова, Парсиева 2010]; гендерну маркованість порівнянь проаналізовано в праці О. Левченко [Левченко 2010]. Окрім того, досліджують **гендерні особливості мови тієї чи іншої мовної особистості** (Ю. Гончар здійснила спробу проаналізувати художній світ Тараса Шевченка в гендерному аспекті [Гончар 2009], Л. Єрмакова – в гендерному аспекті поетичну мову М. Цветаєвої [Ермакова 2014]); Л. Гібадулліна – мовні особливості сучасної жіночої прози [Гібадулліна 2010]; жіночу прозу на матеріалі російської та німецької мов

дослідила З. Хачмафова [Хачмафова 2010]); гендерний дискурс сучасних друкованих україномовних ЗМІ дослідила Ю. Маслова [Маслова 2011]; гендерні стереотипи на матеріалі сучасного британського газетного дискурсу – Т. Лукьянова [Лукьянова 2009]. Сучасні дослідники приділяють увагу **особливостям комунікації в інтернет-середовищі** [Компанцева 2006; Горошко 2009, 2010; Касумова, Мерве Сердар 2013]. У низці праць змодельовано **концепти, пов'язані з маскулінністю/фемінінністю** [Алексеєнко 1999, 2005, 2006, 2010; Kuczyńska, Aleksiejenko 2002; Жеребкин 2006; Ставицька 2006; Мінаєва 2007; Левченко 2008; Рудюк 2010; Сукаленко 2007а, б, в, г, 2008, 2010; Барвіна 2013; Корольова 2014 та ін.].

Під час лінгвогендерних досліджень усе частіше використовують кількісний аналіз. Зокрема, результати досліджень, базовані на аналізі частотності, представлені у працях А. Самохіної [Самохіна 2011; Ю. Павлова, Е. Тихомірової [Павлов, Тихомірова 2011; Мухин 2010]. Ю. Павлов та Е. Тихомірова подають результати дослідження текстів Ф. Достоєвського та М. Горького [Павлов, Тихомірова 2011: 9]: «твердження багатьох авторів, що під час ідентифікації необхідно брати 5-10 тисяч перших слів частотного словника російської мови, можна піддати сумніву, оскільки, як це видно з представлених результатів велике число нерозпізнаних слів якраз лежить у цих межах (принаймні, у авторів, тексти яких статистично опрацювали автори статті)» [Павлов, Тихомірова 2011, 9]. Аналогічної думки дотримується Т. Демидова, яка зазначає, що мовний смак автора, його літературні схильності, багатство мовної особистості відображає саме низькочастотна лексика; і якщо домінантна лексика більшою мірою характеризує менталітет цілого етносу, то одинична максимально індивідуалізована, специфічна [Демидова 2011: 151].

М. Мухін пропонує аналізувати частотну лексику того чи іншого автора для виявлення індивідуально-авторських особливостей стилю, зауважуючи: «Ряди слів, які часто трапляються в текстах різних авторів, природно, мають перетинатися. У центрі схеми – ареал інтегральної, «універсальної» лексики. Цей масив вкрай важливий для побудови будь-якого художнього тексту, але малозмістовний для оцінки індивідуально-авторських особливостей. Відкинувши частотну лексику, характерну для багатьох авторів, можна отримати списки лексем, що становлять індивідуальне слововживання» [Мухин 2010: 61]. Згаданий автор вважає, що: «Показовими для ідіостилю визнано слова, що потрапляють в пік частотності двох і більше творів одного письменника й не більш ніж одного твору іншого автора. Ці критерії прийняті дослідним шляхом, на підставі необхідного обсягу вибірки. Якщо – в ідеалі – відбирати

слова, які часто трапляються в усіх творах автора і зовсім не потрапляють в частотний пік чужих творів, результатом вибірки в кращому випадку буде кілька одиниць. Крім того, увага до слів, які характерні для одного автора й потрапляють в частотний ряд не більш ніж одного чужого твору, пояснюється ідіостилевим, а не монотекстовим підходом» [Мухін 2010: 62].

Н. Беляя стверджує, що «за своїм обсягом «жіночий» словник відносно менший, жінка користується ніби ядром словника (усталеним прошарком лексики), тобто лексичними одиницями з більш високою частотою зустрічальності в мові, тимчасом як чоловік вживає більше неологізмів, професіоналізмів й архаїчних форм слів, не будучи у змозі підшукати їм більш загальноживані слова й вислови» [Беляя 2009: 98]. Однак згадана авторка, на жаль, не уточнює аналіз якого матеріалу і якою мовою дав змогу дійти подібних висновків.

Проте практично недослідженим залишається проблема аналізу колокацій в «жіночих» та «чоловічих» текстах, зокрема художніх. Під колокаціями розуміють «невипадкову сполуку двох та більше лексичних одиниць, характерну для мови загалом (текстів будь-якого типу), так і певного типу текстів (чи навіть (під)вибірки текстів)» [Ягунова, Пивоварова 2010]. За Дж. Р. Фьорсом, колокації – це характерні, часто зустрівані сполуки слів, «поява яких поряд одне з одним полягає на регулярному характері взаємного очікування й задається не граматичними, а чисто семантичними чинниками» [Firth 1957. Цит за: Сусов 2006: 153].

Виокремлюють також колігації. Н. Кочеткова зазначає: «Ще однією статистичною моделлю є колігації. Якщо колокації дають уявлення про поведінку слів і словоформ, то колігації – про поведінку граматичних груп. На відміну від чисто лексичного характеру колокацій, колігації мають змішаний лексико-граматичний характер. Колігація – це комбінація лексичних і граматичних характеристик, які мають тенденцію до сумісної зустріваності» [Кочеткова 2013: 302].

На думку В. П. Захарова, «корпуси й інструментарій корпусної лінгвістики дають змогу виявити й суттєво розширити лексичний фонд стійких словосполук різного типу й особливості їх використання» [Захаров 2015]. В. П. Захаров, М. В. Хохлова стверджують, що найчастіше для виявлення колокацій застосовують такі методи, як MI-score, t-score и log-likelihood [Захаров, Хохлова 2014]. Дослідники зауважують, що найпростіший шлях визначення колокаційної пари спирається на відносну частотність, яка дає найпоширеніші колокаційні асоціації, однак такий метод має недоліки, з огляду на це, одним із варіантів сортування може бути показник VI (показник взаємної

інформації, Mutual Information score, MI), який відображає так зване притягання між двома словами, або ймовірність того, що якщо одне слово з'являється в рядку, слово, яке складає з ним колокацію, з'явиться в найближчій околиці шуканого слова [Горина 2011]. Також застосовують «t-показник, який вимірює вірогідність появи слова як парного в поєднанні з яким-небудь іншим словом. На відміну від показника взаємної інформації (VI), t-показник вимірює тяжіння слова, яке сполучається, з досліджуваним словом, але не навпаки. Таким чином, список таких слів буде містити велику кількість граматичних слів, які важливі для існування досліджуваного слова. З іншого боку, досліджуване слово ніяк не впливає на існування граматичних слів. Жоден з перерахованих вище способів і показників не можна виокремити як найкращий у визначенні колокацій» [Горина 2011]. Е. Ягунова, Л. Пивоварова дійшли висновку про те, що принципово відрізняються списки колокацій, отримані за допомогою VI та t-score: VI найкраще дає змогу виокремлювати назви об'єктів, терміни, складні номінації; t-score, навпаки, краще працює під час виокремлення «загальнономовних стійких сполук» (похідних службових слів, дискурсивних слів) й «стійких конструкцій» [Ягунова, Пивоварова 2010].

У праці Т. Бобкової розглянуто колокації як маркери функціональних стилів [Бобкова 2015]. Згадане дослідження виконано на основі вибірки з Корпусу текстів української мови обсягом 1 млн слововживань, проаналізовано тексти офіційно-ділового і публіцистичного стилів [Бобкова 2015: 39]. Т. Бобкова зазначає, що «істотне переважання загальної кількості колокацій у наукових текстах досягається в основному за рахунок низькочастотних сполук, які становлять периферію лексику аналізованих текстів: 52,8% сформованого списку становлять стійкі сполуки з мінімальною частотою (2 рази на 1 млн) проти 50,5% – у підкорпусі публіцистики та 42,9% – у законодавчих текстах» [Бобкова 2015: 39, 40].

**Мета цього дослідження** – на основі створеного дослідницького корпусу виявити колокації, наявні в українській художній прозі різних часових періодів та дослідити особливості вживання колокацій у художніх текстах чоловіків та жінок.

Для проведення аналізу було створено дослідницькі підкорпуси. Підкорпус текстів 1961-1990 рр., який містить тексти (обсяг становить приблизно 710 тис. слововживань)) **письменниць** (І. Вільде «Сестри Річинські» (частина II), Г. Пагутяк «Господар», «Філософський камінь», Н. Бічужа «Біла Віла», «Килим на три квітки», «Чистотіл», Г. Гордасевич (оповідання «Романс Рахманінова», «Твій тихий дім», «Командувати парадом буду я», «Сісти і подумати», «Дорога додому», «Дощ у

вихідний», «Весна 46-го року», «Зіпсований годинник», «Транзитні пасажири», «Життя на пенсії», «Відцвіла шипшина», «Клопіт з рибою», «При виконанні бойового завдання», «Чому дерева зелені», «Рита», «Втеча на рожеву планету», «Мандрівка в минуле», «Живе собі хлопчик»); К. Мотрич «Молитва до мови») та **письменників** (Гр. Тютюнник «Вир», П. Загребельний «З погляду вічності»; «Безслідний Лукас», О. Гончар «Циклон», «Берег любові», В. Шевчук «Дім на горі», Є. Гуцало «Мертва зона», «Голодомор», «Княжа гора»). О. Стяжкіна наводить цікаві дані про те, що у «збірці «Преса Української РСР. 1918–1985», у таблиці «Видання в Українській РСР творів окремих українських письменників. Радянський період» серед сорока одного представника радянської літератури була названа лише одна жінка – Ірина Вільде. Загальний наклад видань її творів склав 1 346,7 тис. примірників, це було менше, ніж загальний наклад П. Загребельного (5 492,4 тис.), Ю. Збанацького (5 938,2 тис.), О. Вишні (8 133,8 тис.), П. Воронька (6 575,1 тис.), але, в цілому, дорівнював середньостатистичному показнику» [Стяжкіна 2015].

Підкорпус текстів 2000-х рр. містить уривки з прози (обсяг приблизно 927 тис. слововживань): **письменниць** (Люко Дашвар «Біті є. Макар», О. Забужко «Музей покинутих секретів», Л. Костенко «Записки українського самашедшого», І. Роздобудько «Ескорт у смерть», «Зів'ялі квіти викидають», «Гра в пацьорки», Є. Кононенко «Зрада ZRADA made in Ukraine») та **письменників** (Ю. Андруховича «Дванадцять обручів», Ю. Винничука «Танго смерті», «Весняні ігри в осінніх садах», Л. Дереша «Трохи п'єми», «Голова Якова», С. Жадана «Біг Мак та інші історії: книга вибраних оповідань», «Депеш Мод», В. Шкляра «Елементал»).

На матеріалі дослідницьких корпусів здійснено спробу опрацювання методики аналізу n-грам. Для аналізу застосовано алгоритм "Суфіксне дерево" розподілу тексту на кластери, або так звані n-грами, колокації. У галузі опрацювання природної мови, n-грами використовують переважно для передбачення на основі ймовірнісних моделей. N-грамна модель розраховує ймовірність останнього слова n-грами, якщо відомі всі попередні. Під час використання цього підходу для моделювання мови передбачають, що поява кожного слова залежить тільки від попередніх слів. У цьому дослідженні загальний масив текстів було розподілено на 2, 3, 4, 5, 6-грами. Також було зведено, відсортовано та підраховано кількісні показники кластерів. Підраховано також довжину речень у кожному з підкорпусів.

Результати підрахунків засвідчують той факт, що «найкоротші» речення наявні в досліджених творах сучасних українських письменниць, натомість найдовші

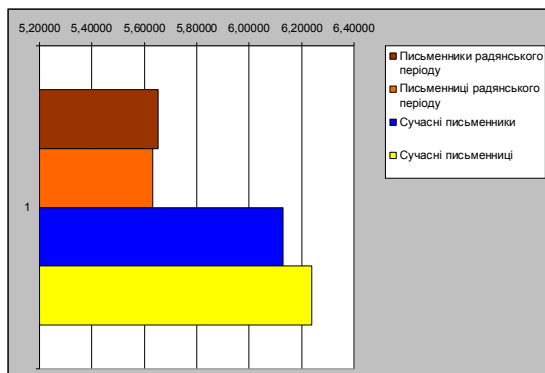
спостерігаємо в текстах українських письменників радянського періоду (див. Таблиця 1. *Середня довжина речення та слова в досліджуваних текстах*; Діаграма 1. *Різноманітність лексики*; Діаграму 2. *Середня довжина речення (кількість слів)*).

Таблиця 1. *Середня довжина речення в досліджуваних текстах*

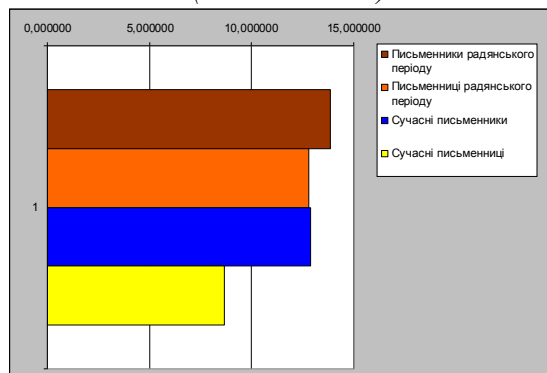
Підкорпуси	Різноманітність лексики	Середня довжина речення (кількість знаків)	Середня довжина речення (кількість слів)	Середня довжина слова
Письменники радянського періоду	5,652	84,772	13,851	5,062
Письменниці радянського періоду	5,631	76,377	12,793	4,882
Сучасні українські письменники	6,130	80,455	12,888	5,022
Сучасні українські письменниці	6,238	54,468	8,672	5,135

Із даних, отриманих на матеріалі цих підкорпусів, випливає, що найбільша різноманітність лексики властива сучасним українським письменницям. Хоча загалом показники різноманітності лексики в межах досліджуваних періодів відрізняються незначно.

Діаграма 1. *Різноманітність лексики*



Діаграма 2. *Середня довжина речення (кількість слів)*



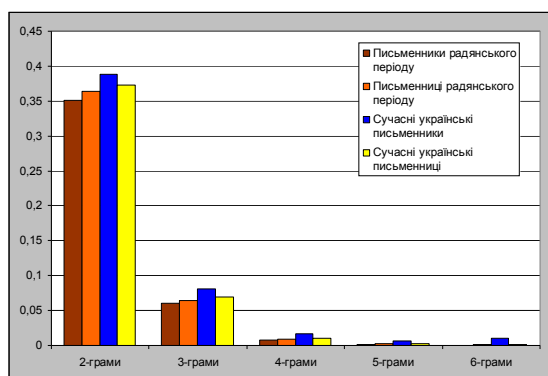
У Таблиці 2. *Відносна частотність n-грам у підкорпусах* представлено дані кількісного аналізу n-грам у досліджуваних текстах. Як свідчать здобуті результати, найвища частотність властива 2-грамам.

Таблиця 2. Відносна частотність n-грам у підкорпусах

n-грами	Письменники радянського періоду	Письменниці радянського періоду	Сучасні українські письменники	Сучасні українські письменниці
2-грами	0,351225358	0,363455988	0,388194632	0,372500928
3-грами	0,060363640	0,063668320	0,081445334	0,069662702
4-грами	0,007861287	0,008764632	0,016233801	0,010835238
5-грами	0,001784221	0,001952911	0,006921252	0,003190672
6-грами	0	0,000847567	0,010459668	0,001764641

Ці ж результати продемонстровано на Діаграмі 3. *Відносна частотність n-грам у підкорпусах*. З аналізу випливає, що найбільше 2-грам спостерігаємо в текстах сучасних українських письменників, а найменше – письменників радянського періоду. Найбільшу кількість 3- і 4-грам містять твори сучасних українських письменників та письменниць.

Діаграма 3. Відносна частотність n-грам у підкорпусах



Окрім того, проанговано частотності 2-грам (результати подано у Таблиці 3. *Розподіл за рангами частотності 2-грам*).

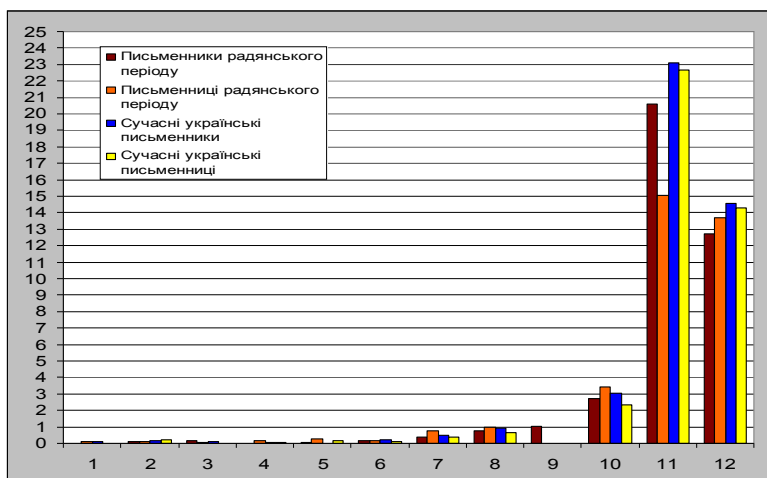
Таблиця 3. Розподіл за рангами частотності 2-грам

Ранги	Відносна частотність одиниць			
	Письменники радянського періоду	Письменниці радянського періоду	Сучасні українські письменники	Сучасні українські письменниці
0,001-0,009	0	0,0011171	0,0013263	0
0,0009-0,0001	0,0546912	0,0761080	0,0607204	0,0525741
0,00009	0,0038971	0,0029723	0,0065302	0,0051506
0,00008	0,0067657	0,0063195	0,0041951	0,0065701
0,00007	0,0052923	0,0095106	0,0079948	0,0058387
0,00006	0,0093146	0,0081396	0,0088308	0,0089056

	Відносна частотність одиниць			
0,00005	0,0082257	0,0137718	0,0097338	0,0083890
0,00004	0,0147678	0,0124360	0,0144516	0,0137632
0,00003	0,0154766	0,0261882	0,0203057	0,0204527
0,00002	0,0308860	0,0260750	0,0243970	0,0251193
0,00001	0,0506510	0,0879620	0,0652114	0,0641408
0,000009	0,1512576	0,0928559	0,1644975	0,1615967

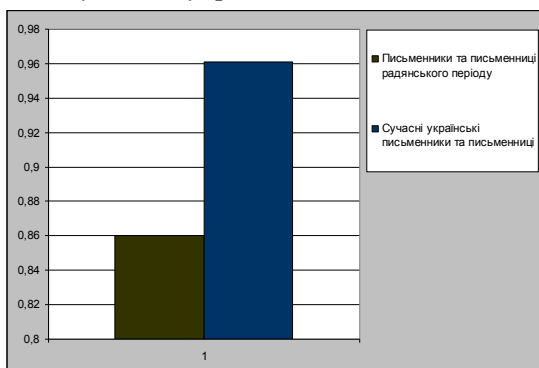
Розподіл за рангами частотності 2-грам подано на Діаграмі 4.

Діаграма 4. Розподіл за рангами частотності 2-грам

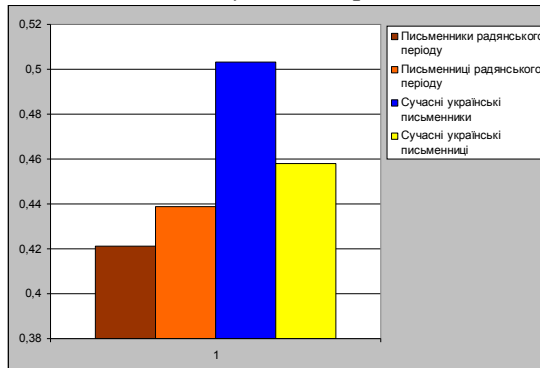


Попередні результати аналізу засвідчують той факт, що сучасним художнім текстам властива вища частотність колокацій, «чоловічі» тексти досліджених періодів містять більше колокацій, аніж «жіночі» (див. Діаграма 5. Відносна частотність  $n$ -грам у текстах радянського періоду та сучасних українських текстах, Діаграма 6. Відносна частотність  $n$ -грам у текстах письменників та письменниць досліджуваних періодів).

Діаграма 5. Відносна частотність  $n$ -грам у текстах радянського періоду та сучасних українських текстах



Діаграма 6. Відносна частотність  $n$ -грам у текстах письменників та письменниць досліджуваних періодів





Сформовано загальний список найчастотніших сполук, виокремлених із кожного із підкорпусів (див. Таблиця 4. *Найчастотніші 2-грами, виокремлені з підкорпусів*).

Таблиця 4. *Найчастотніші 2-грами, виокремлені з підкорпусів*

2-грама	Відносна частотність			
	Письменники радянського періоду	Письменниці радянського періоду	Сучасні українські письменники	Сучасні українські письменниці
і не	0,0009368	0,0009804	0,0009764	0,0009789
й не	0,0008116	0,0006601	0,0004774	0,0005860
і в	0,0008094	0,0004414	0,0003867	0,0003798
те що	0,0006104	0,0007694	0,0007517	0,0007163
не було	0,0005478	0,0007538	0,0005508	0,0005925
містер Ор	0,0005411	0,0000000	0,0000000	0,0000000
і він	0,0005210	0,0002148	0,0003046	0,0002691
на нього	0,0004829	0,0003593	0,0002311	0,0002019
до нього	0,0004762	0,0002812	0,0002333	0,0002431
та й	0,0004583	0,0004023	0,0001944	0,0002887
що він	0,0004338	0,0005937	0,0004558	0,0003798
що в	0,0004293	0,0005312	0,0003953	0,0003950
я не	0,0004181	0,0011171	0,0013263	0,0009594
ще й	0,0004159	0,0002890	0,0003499	0,0003603
це не	0,0004148	0,0003359	0,0004148	0,0002995
так і	0,0004136	0,0003750	0,0004126	0,0004189
так само	0,0004114	0,0002265	0,0003759	0,0004059
чи не	0,0004025	0,0003515	0,0002376	0,0002127
не міг	0,0004002	0,0003632	0,0003197	0,0001997
а в	0,0003980	0,0002773	0,0002527	0,0002539
він не	0,0003779	0,0004257	0,0003046	0,0003321
що не	0,0003689	0,0006562	0,0003564	0,0003798
що ж	0,0003689	0,0003047	0,0001728	0,0003017
а ти	0,0003667	0,0004296	0,0002484	0,0003039
а коли	0,0003644	0,0002461	0,0002765	0,0001346
нічого не	0,0003555	0,0006445	0,0004860	0,0004601
в нього	0,0003533	0,0002695	0,0002592	0,0001671
а я	0,0003466	0,0006132	0,0005249	0,0005752
ніколи не	0,0003398	0,0004609	0,0004040	0,0006663
ще не	0,0003398	0,0004648	0,0003219	0,0003429
як і	0,0003398	0,0002617	0,0003262	0,0003017
до неї	0,0003354	0,0003671	0,0002095	0,0002192
а що	0,0003331	0,0004921	0,0002895	0,0003907
вже не	0,0003197	0,0004179	0,0003024	0,0004015
але ж	0,0003153	0,0003125	0,0001599	0,0002171
і на	0,0003130	0,0002070	0,0001966	0,0002301
на неї	0,0003108	0,0002695	0,0001793	0,0002518
так що	0,0003108	0,0002656	0,0002095	0,0002214
що це	0,0003041	0,0004218	0,0004515	0,0004189

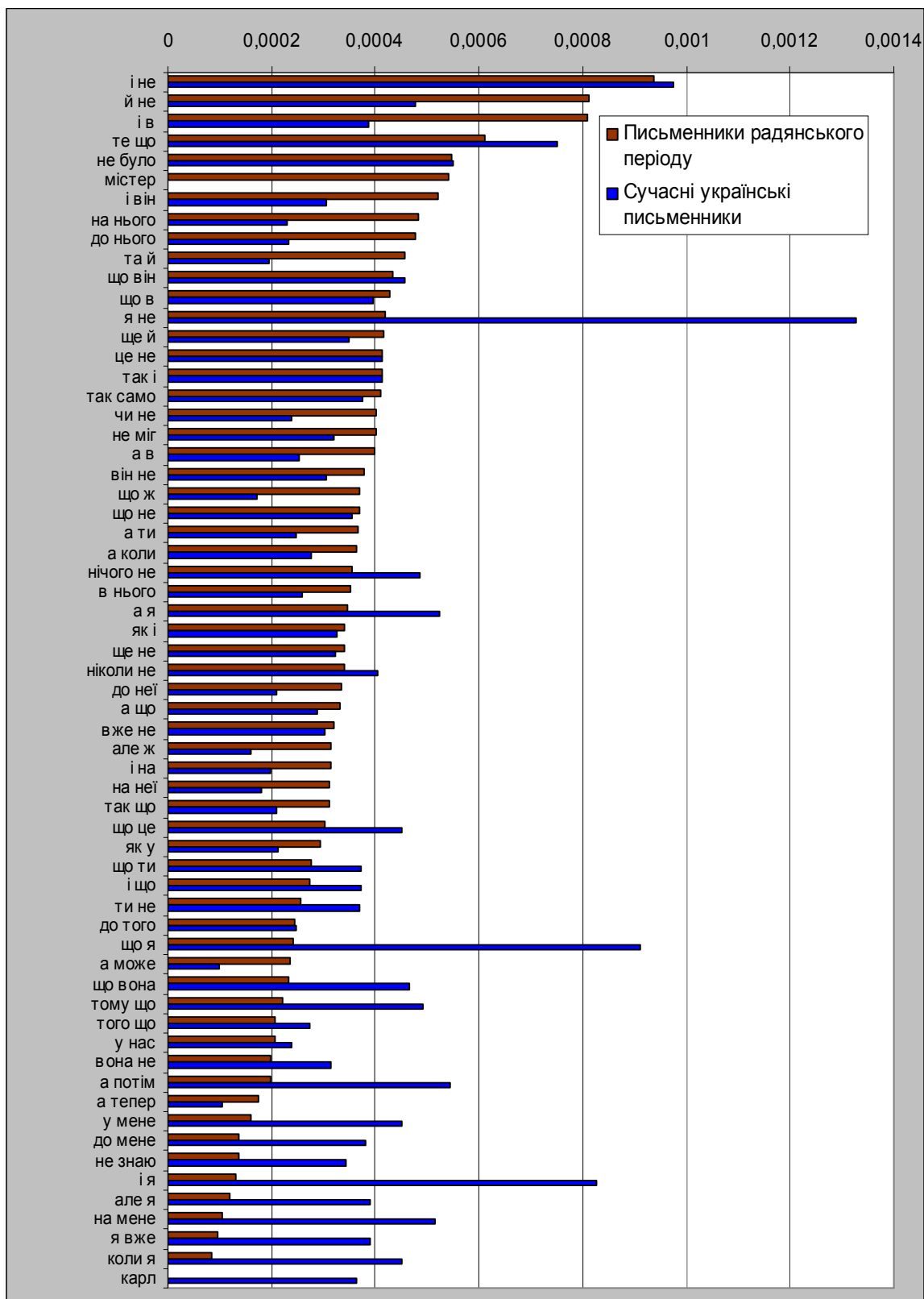
як у	0,0002951	0,0002226	0,0002139	0,0004515
що ти	0,0002772	0,0005234	0,0003737	0,0003994
і що	0,0002728	0,0002851	0,0003737	0,0003581
ти не	0,0002549	0,0003554	0,0003694	0,0002735
до того	0,0002437	0,0004453	0,0002463	0,0002257
що я	0,0002415	0,0008085	0,0009116	0,0006034
а може	0,0002348	0,0004101	0,0000994	0,0001129
що вона	0,0002325	0,0004960	0,0004666	0,0004232
тому що	0,0002213	0,0004296	0,0004925	0,0003972
того що	0,0002079	0,0004726	0,0002743	0,0001910
у нас	0,0002057	0,0001406	0,0002398	0,0005209
а потім	0,0001968	0,0004218	0,0005444	0,0004298
вона не	0,0001968	0,0004218	0,0003154	0,0003777
а тепер	0,0001744	0,0004218	0,0001037	0,0001975
у мене	0,0001587	0,0001484	0,0004515	0,0006034
не знаю	0,0001364	0,0004296	0,0003435	0,0003321
до мене	0,0001364	0,0003203	0,0003824	0,0002670
і я	0,0001297	0,0003281	0,0008252	0,0004840
але я	0,0001185	0,0002461	0,0003910	0,0002105
на мене	0,0001051	0,0002851	0,0005141	0,0002865
я вже	0,0000961	0,0003203	0,0003888	0,0002757
коли я	0,0000850	0,0001601	0,0004515	0,0002084
Карл Йозеф	0,0000000	0,0000000	0,0003651	0,0000000

Переліки, сформовані за частотністю, очолюють прийменниково-займенникові сполуки, займенники/прийменники з часткою не, складні сполучники, дієслова з часткою не, дієслова із прийменниками. З огляду на невеликий обсяг дослідженого матеріалу, до найчастотніших 2-грам потрапила низка онімів (імен героїв художніх творів), як-от: *містер Ор*, *Карл Йозеф*.

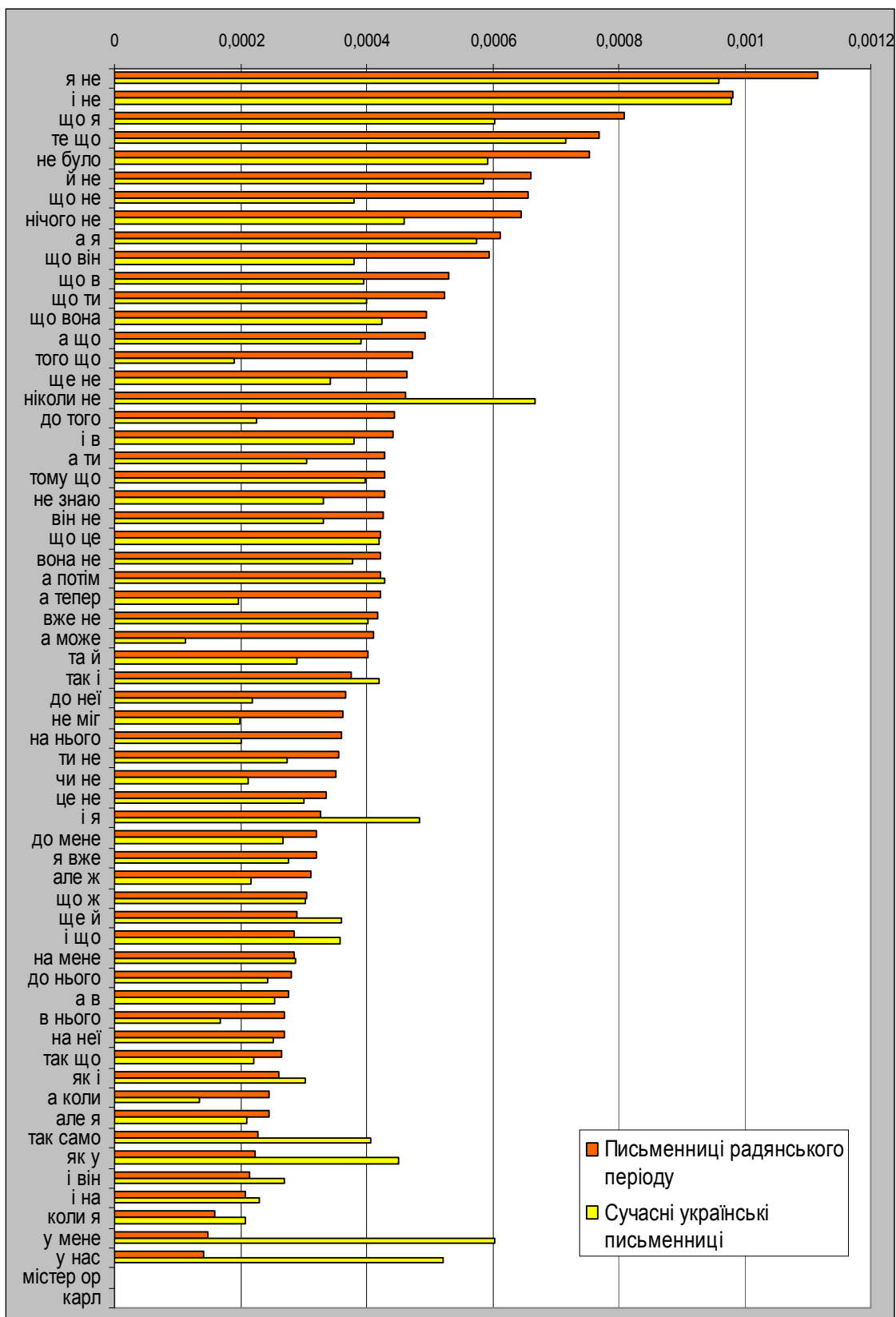
На діаграмах подано зіставлення найчастотніших сполук у творах письменників радянського та сучасного періодів (Діаграма 7. *Зіставлення відносної частотності сполук у творах радянських і сучасних письменників*) та письменниць радянського та сучасного періодів (Діаграма 8. *Зіставлення відносної частотності сполук у творах радянських і сучасних письменниць*).

У творах письменниць радянського періоду найвищу частотність мають 2-грами: *я не*, *і не*, *що я*, *те що*, *не було*, *й не*, *що не*, *нічого не*, *а я*, *що він*, *що в та ін.* У творах сучасних письменників найчастотнішими виявилися: *я не*, *і не*, *що я*, *і я*, *те що*, *не було*, *а потім*, *а я*, *на мене*, *тому що*, *нічого не та ін.*; письменниць: *і не*, *я не*, *те що*, *ніколи не*, *що я*, *у мене*, *не було*, *й не*, *а я*, *у нас та ін.* Загалом найвища відносна частотність леми *я* властива текстам сучасних українських письменників (див. Діаграму 9. *Відносна частотність леми я*).

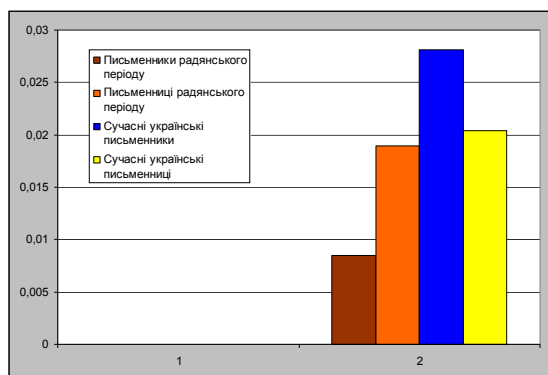
Діаграма 7. Зіставлення відносної частотності сполук у творах радянських і сучасних письменників



Діаграма 8. Зіставлення відносної частотності  
сполук у творах радянських і сучасних письменниць

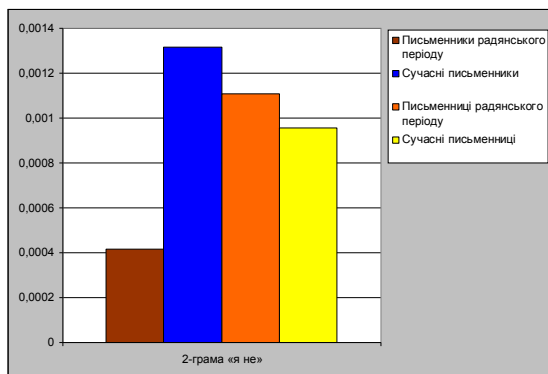


Діаграма 9. Відносна частотність леми **я**

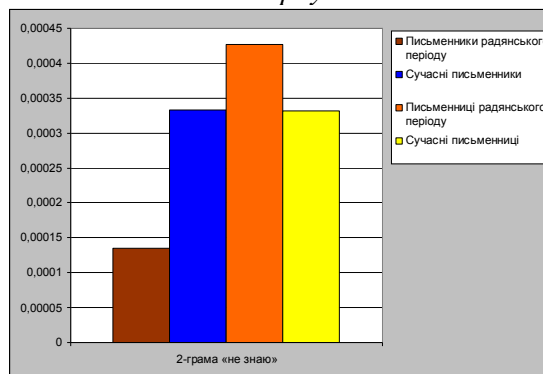


Звертає на себе увагу висока частотність вживання поряд слів **Я** і **НЕ** у творах сучасних українських письменників. На Діаграмі 10. *Відносна частотність 2-грами **я не** в досліджуваних підкорпусах* подано зіставлення даних. На Діаграмі 11. *Відносна частотність 2-грами **не знаю** в досліджуваних підкорпусах* подано зіставлення даних про уживання 2-грами *не знаю*: найвища частотність уживання цієї 2-грами властива творам радянських письменниць.

Діаграма 10. Відносна частотність 2-грами **я не** в досліджуваних підкорпусах



Діаграма 11. Відносна частотність 2-грами **не знаю** в досліджуваних підкорпусах



Визначено також коефіцієнт взаємної інформації для виявлених 2-грам. Коефіцієнт МІ обраховують за формулою:

$$VI = \log^2 \frac{f(n,c) \times N}{f(n) \times f(c)},$$

де  $n$  – ключове слово,  $c$  – колокат,  $f(n, c)$  – частота зустрітваності ключового слова в парі з колокатом,  $f(n)$ ,  $f(c)$  – абсолютні частоти  $n$  і  $c$  у корпусі,  $N$  – загальна кількість словоформ у корпусі [Шишигіна 2012]. О. Шишигіна зазначає: «Чим вище коефіцієнт

MI, тим менше випадковою виявляється ймовірність зустрічі двох слів у складі словосполуки. Уважають, що сполука слів є статистично значущою, якщо значення MI більше 1. Однак, у корпусі СОСА, на матеріалі якого проведено дослідження, зазначено, що про існування семантичного зв'язку між словами можна говорити тільки в тому випадку, якщо коефіцієнт взаємної інформації між ними дорівнює, як мінімум, 3. Враховуючи це твердження, а також те, що максимальне значення MI, зафіксоване в словосполучках з лексемою тап, дорівнювало 10,0, низьким значенням MI в цій праці вважатимемо діапазон 1,0–2,9, середнім 3,0–5,0 і високим 5,1 і вище [Шишигина 2012].

У Таблиці 5. *Коефіцієнт взаємної інформації 30 найчастотніших 2-грам* представлено дані про семантичний зв'язок між компонентами 2-грам. З огляду на незначний обсяг досліджених текстів, високим вважатимемо коефіцієнт 29-21, середнім – 20-11, низьким – 10-1.

Таблиця 5. *Коефіцієнт взаємної інформації 30 найчастотніших 2-грам*

2-грама	Коефіцієнт VI			
	Письменники радянського періоду	Письменниці радянського періоду	Сучасні українські письменники	Сучасні українські письменниці
містер Ор	9,538481784	0	0	0
так само	5,367256945	4,591013253	4,9936526830	5,109512653
те що	2,158408654	1,897081609	2,3986568172	2,176050819
ніколи не	2,106587451	1,685799742	2,0621081662	2,082416628
не міг	2,104821322	1,389799627	1,6410192920	1,510639342
у нас	1,915247198	1,402173882	1,3065223617	2,299643507
до нього	1,888791973	1,547238191	1,5722432592	1,639306874
нічого не	1,730649677	1,390646369	1,5674994138	1,585121698
не знаю	1,64944456	1,683674565	1,8644674188	1,675301007
а потім	1,635283357	2,008934012	2,4966928232	2,365933479
до того	1,602002703	1,73028346	1,6610714329	1,73028346
тому що	1,334976214	1,311381993	1,6965783358	1,560112692
а тепер	1,133876529	1,329646479	0,8816825815	1,111774726
на нього	1,041880224	1,181626181	0,8092933201	0,791709919
у мене	0,988547599	0,405147077	0,7965472604	1,172067975
а може	0,940540109	1,000956629	0,7907450089	0,829011544
до мене	0,92685346	0,9264981913	0,9264981913	0,946294507
того що	0,920306777	1,023913093	1,1606626333	0,92034933
а коли	0,891423263	0,429314108	0,7896775446	0,437146865
не було	0,888861446	0,673264483	0,8214877948	0,919137976
ще й	0,88561491	0,953887066	1,2240459916	1,168703952
в нього	0,854321955	1,055428578	1,1932385081	0,91213507
а ти	0,843399437	0,668590368	0,5187485399	0,63640844
та й	0,802171055	1,383598232	1,383598232	0,849832149
чи не	0,66246855	0,267021281	0,4877355321	0,350241606
що ж	0,624156085	0,420806607	0,4820096433	0,574559986
я вже	0,578339824	0,80782703	0,7541027804	0,668660689

а я	0,562463212	0,366688708	0,1895669406	0,395941209
але я	0,528036495	0,335528688	0,2802695971	0,450106069
що ти	0,485807046	0,552315684	0,5418337671	0,588368435
вже не	0,451969284	0,386573516	0,4921276534	0,535992302
коли я	0,435808831	0,293349211	0,7107623883	0,54925748
як у	0,385684009	0,163532232	0,1896219280	0,333596702
що це	0,385251025	0,36136783	0,4010291703	0,45374042
ще не	0,373055777	0,368685801	0,3747025824	0,40560401
я не	0,324703827	0,298505274	0,2721912600	0,242569922
що він	0,32031212	0,445600189	0,3844023791	0,420890807
що вона	0,306869024	0,436198584	0,4985536386	0,403477215
так і	0,271127809	0,163264951	0,1797739509	0,284703092
ти не	0,248772834	0,172776704	0,3321272663	0,154522513
що я	0,24316781	0,31969134	0,2637873814	0,250647812
й не	0,240535552	0,251058243	0,2286015199	0,204512155
на мене	0,239921891	0,434249872	0,5489439183	0,344458683
це не	0,131669615	0,11748552	0,1950854868	0,101387067
і він	0,129053701	0,011508694	0,0188361135	0,082520257
він не	0,118567122	0,132251474	0,0842159360	0,14508738
вона не	0,102205848	0,185434864	0,1453467268	0,141528106
а що	0,080924854	0,076651174	0,0750221954	0,177920641
а в	0,077742815	0,025308972	0,0707946444	0,068819849
що в	0,04484152	0,079052112	0,0893060509	0,092921804
і в	0,040357005	0,006712226	0,0003404296	0,005763272
як і	0,037249041	0,000821137	0,0380038571	0,019060002
і що	0,035560728	0,057201359	0,0071144888	0,000517403
і не	0,034435568	0,018969812	0,0316861886	0,062464059
що не	0,004504172	0,006746968	0,0023083712	0,002567534
і я	0,003987955	0,003019288	0,0264580851	0,037259129
Карл Йозеф	0	0	11,7821688442	0

У Таблиці 6 представлено дані про відносну частотність 2-грам, які мають коефіцієнт ВІ від 29 до 1.

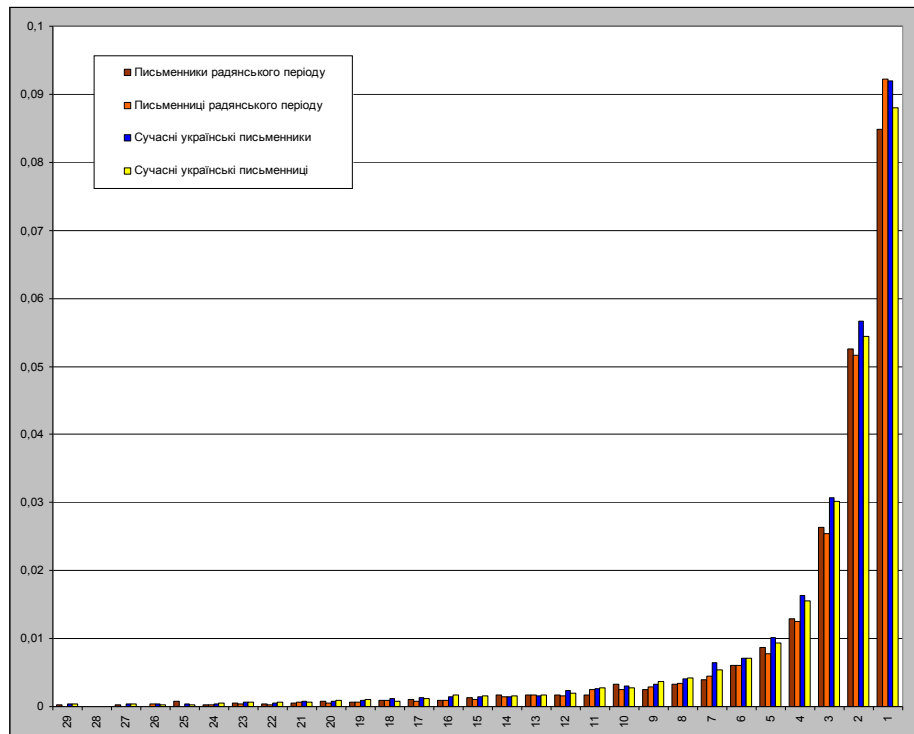
Таблиця 6. Відносна частотність 2-грам, які мають коефіцієнт ВІ від 29 до 1

Коефіцієнт ВІ	Відносна частотність			
	Письменники радянського періоду	Письменниці радянського періоду	Сучасні українські письменники	Сучасні українські письменниці
29	0,0003130177	0	0,0003629082	0,0004601463
28	0	0	0	0
27	0,0002861876	0	0,0004406742	0,0003711558
26	0	0,0003358940	0,0003672285	0,0002908472
25	0,0007288842	0	0,0004536352	0,0002973587
24	0,0002392350	0,0003007425	0,0003693887	0,0004709988
23	0,0004784700	0,0003710459	0,0006221283	0,0005990584
22	0,0004583474	0,0002109103	0,0005141199	0,0006424685
21	0,0005142434	0,0006366367	0,0007646994	0,0006489800
20	0,0007646576	0,0005389930	0,0007884612	0,0008660301
19	0,0007109974	0,0006717884	0,0008943095	0,0010874213

18	0,0009547041	0,0008866045	0,0012161745	0,0008486661
17	0,0010329585	0,0007303750	0,0012852998	0,0011460248
16	0,0009860059	0,0009725300	0,0013911481	0,0016756272
15	0,0013124387	0,0011092300	0,0014473124	0,0015779547
14	0,0017014750	0,0014451260	0,0014429921	0,0015844662
13	0,0017148900	0,0017458690	0,0015553208	0,0017060142
12	0,0017238334	0,0015505814	0,0023502626	0,0019404284
11	0,0017238334	0,0024449973	0,0026505259	0,0027131270
10	0,0033559973	0,0025035835	0,0029723909	0,0027890945
9	0,0024929627	0,0028394777	0,0032337712	0,0036421017
8	0,0032419694	0,0034448684	0,0040848773	0,0041717040
7	0,0039708536	0,0045111372	0,0064805034	0,0054197424
6	0,0060859591	0,0060773416	0,0070637487	0,0071235862
5	0,0086772988	0,0078349275	0,0100987845	0,0094112949
4	0,0129589343	0,0125218234	0,0163567906	0,0155082338
3	0,0263158483	0,0254615616	0,0306463006	0,0301526080
2	0,0525243763	0,0516691208	0,0566460803	0,0543840876
1	0,0847942691	0,0922849789	0,0919194603	0,0880246916

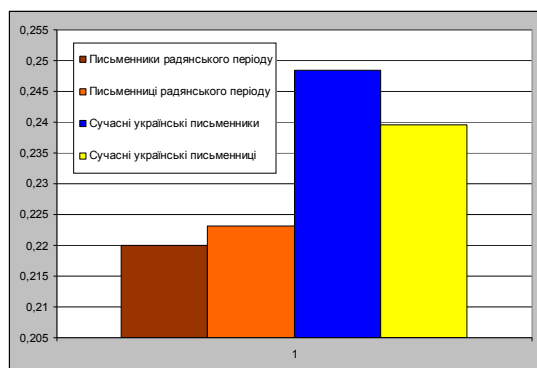
Як засвідчують отримані дані, найнижча відносна частотність властива 2-грамам із найвищим коефіцієнтом ВІ (див. Діаграма 12. *Відносна частотність 2-грам, які мають коефіцієнт МІ від 29 до 1*). На Діаграмі 13 унаочнено дані про кількість 2-грам у досліджених підкорпусах. Найбільша кількість 2-грам, коефіцієнт ВІ яких міститься в діапазоні 29-1, наявна у творах сучасних українських письменників.

Діаграма 12. *Відносна частотність 2-грам, які мають коефіцієнт МІ від 29 до 1*





Діаграма 13. Відносна частотність 2-грам з показником ВІ від 29 до 1 в досліджених підкорпусах



У Таблиці 7 подано приклади колокацій із коефіцієнтом ВІ від 29 до 1, докладний аналіз формальної будови яких заплановано здійснити в подальшому. Однак, зауважимо, що високий коефіцієнт ВІ мають 2-грами формальної будови *прикметник + іменник, іменник + іменник*. Високий коефіцієнт ВІ властивий онімам, фразеологізмам, жаргонізмам, терміносполукам, а також парам слів, які введено в текст іншими мовами, зокрема російською, як-от: *Чарльз Спенсер, Миколою Миколайовичем, буриданів віслук; експериментальних господарств, сечовий міхур, этнографическая протоплазма, апасний штріх*. Для дієслівних сполук та сполук із займенниками здебільшого характерним є низький коефіцієнт ВІ: *намагалася уявити, майнуло щось, пропонував усім, пішли ми; любили ми; бачив тисячі; дуже хотілося; зостався сам*.

Таблиця 7. Колокації із коефіцієнтом ВІ від 29 до 1

ВІ	Письменники радянського періоду	Сучасні українські письменники	Письменниці радянського періоду	Сучасні українські письменниці
	Приклади 2-грам			
29	всенародні ілюзії; Чарльз Спенсер; ташанська русалка; случаю випечки; прикордонною вишкою	череві вантажного; покликано телеграмою; лечо Гелісхана; алу парата; буриданів віслук	0	слухняну ученицю; сутяжнопараноїдальна психопатія; Алексей Васильч; королівстві Непал; Садама Хусейна
28	0	0	0	0
27	білків жирів; Рекел Велч; исправительной конюшни; “абсолютне ніщо”; “місіс Касандра”	писаної скали; золотіла підсмажена; апасний штріх; карпатська ініціатива; тісним колом	0	Миколою Миколайовичем; трамвайний пристанок; душова кабінка; преподобної Горгонії; сміттєвому контейнері
26	0	макове зерня; рибка цитринка;	матко боска; гетьманських валах; рожеву	унікальний мислительний; сечовий міхур;

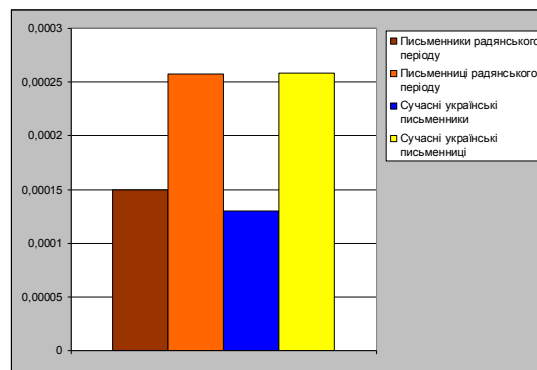
		експериментальних господарств; Волт Вітмен; чинхойського тейпу	планету; качачу лапку; Гандзю Книш	солдатське спідне; Федір Олександрович; племені майя
25	крученими паничами; Олександрє Панасівно; Галино Іванівно; Радянського Союзу; Макса планка	секретар райкому; запорошеними курявою; чоловічим відблиском; Натан Гозендуфт; геєні огненній	0	сталева пружинка; вогняний контур; добору кадрів; цокплюсь цокплюсь; доброму гуморі
24	сія сія; розумового потенціалу; університетської ліги; куряче яєчко; етнографіческа протоплазма	скурвий сину; дядечко Толя; тьотя Настя; й*чмо; лісових бобрів	св. Мартіна; психічної структури; дяді Васі; Юліє Вікентіївно; весняної зливи	підземному переході; нерозбірливо бура; токо моно; Амброза Бірса; підземного переходу
23	золотої орди; розумовий потенціал; Ягуара Ягуаровича; східному узбережжі; Варвари Пилипівни	плавали вушка; грудній клітці; ласкаво просимо; технічний спирт; козацьким борщем	національних меншостей; Анна Пилипівна; Радянського Союзу; сімнадцятий об'єкт; командувати парадом	московський патріархат; сміттєвих баків; Латинську Америку; сумочки знайденої; Андріївну Вербицьку
22	порожніх класах; вай вай; красивий доцент; знаменитий мовознавець; колючих дротів	запалення ясен; дул ветер; нікуди нікуди; Янівського цвинтаря; епохальний твір	фонд праці; Юрка Зарічного; Антоніна Петрівна; нотаток пенсіонера; старими вітьряками	Альфреда Вікторовича; розтрочене дитяче; вигнула брову; нервовий зрив; точна копія
21	сторожа Степана; запасні частини; сільську околицю; скіфів знаменитий; шкіряного капшука	дель Кампо; протипожежної безпеки; мирно гомонять; задзвеніли шиби; поховального закладу	Христина Павлівна; Орестові Білінському; Радянський Союз; угодову політику; показала Ількові	місячне сяйво; фея фуршетів; Мартиними апартаментами; Лариса Дмитрівна; вагомий аргумент
20	глузду з'їхав; народної освіти; московський уряд; Миколи Хашуватого; двоповерховому котеджі	Варцабич Ілько; Бальзаму Варцабича; Робін Гуда; здоровий глузд; Робін Гуд	службових обов'язків; точки зору; докторові Гукові; Михайлович Процик; важкого ушкодження	шістнадцятого травня; Ляля Броніславівна; к чорту; Оболонській набережній; земної кулі
19	соціальної справедливості; хромові чобітки; радянської влади; Павлові Музиці; зводив сака	поцілунок смерти; жодним порухом; сирі яйця; обмовився жодним; протилежній стіні	тисяча дев'ятсот; привокзальної площі; Землях роменських; закопилила губу; захопити владу	Альфред Вікторович; закусила губку; Романа Олексійовича; кандидат наук; чудовисько Амброза
18	конячий лазарет; високих каблуках; впоперек каменя; дерев'яні орли; кактуса колючого	Звіада Гамсахурдіа; рада директорів; повній програмі; Холодного Яру; пральна машина	світова війна; сучасну молодь; панові меценасові; служіння ідеї; посмарованою маслом	йдеш идеш; амулет Паскаля; сіра миша; вісімдесят п'ятий; перших порах
17	пана Волинського; золота орда; засідання правління; працівники ферми; лункий лід	спільної мови; паспорт моряка; маки рікою; Кортумову гору; гнилі ясна	Ореста Білінського; кожним прибором; твої совети; слідчий суддя; стрійка Нестора	біжи біжи; великою актрисою; білій сорочці; довгі гудки; бальзаківського віку
16	Маріїної сестри; доктором Ором; раннього літа;	Високий Замок; психіатричної лікарні; геніальна ідея;	т. п.; прихильників освіти; начальник техконтролю; свекруха	Іветто Андріївно; Ваня Корбут; дитячий садок; чергову порцію;

	підіймаються знизу; бабу Явдоху	нічному клубі; любові любові	лиха; будь-кого товариства	чорним королем
15	купчасті хмари; ванній кімнаті; шостий флот; вічна молодість; подертій сорочці	кожному кроці; нестримна сила; повітряний поцілунок; Зельмана Мількера; самотній ворон	Кость Філіпчук; кожному кроці; півтори години; ушкодження тіла; спинив машину	босу ногу; білий крокус; Іветти Андріївни; віддати найдорожче; різні інстанції
14	шкільний двір; глухий гомін; сонним селом; далекий спогад; голуба богиня	закохуюємося іноді; картку паперу; перестав їздити; доти доки; сучий син	всякому разі; пішли снідати; цієї ночі; розуму сходжу; кров зобов'язувала	адвокат Поспелов; темній воді; остаточну відповідь; ся маєте; втратила дар
13	зелених трав; моя візитка; вдома пани; Герцен пише; державної комісії	лягав спати; шукає метал; довгими вусами; візьми консерви; поверхом вище	нервовими руками; Тарас Чиж; голова розболілася; людського духу; знизу вгору	пологового будинку; всьому місту; задзвонив телефон; мелеш Лідусю; тітка Світлана
12	місяць уповні; жилах тече; запаху диму; Бєєвої гори; ледь помітно	місцевого пива; вдягнеш ту; ідея сподобалася; порошок пральний; догори брови	п'ятнадцять злотих; вулиці Легіонів; концерт Мединського; пам'ять покійного; цих прикмет	шлюбної ночі; пізнього вечора; дурєпа Люсі; темні постаті; різними людьми
11	білій сорочці; терор знизу; ваша згода; ніяких листів; незвична тиша	чесно кажучи; праворуч підеш; брати участь; великий секрет; півтори години	боком лица; міцно тримав; рідній школі; нова хвиля; вигадуєш мамо	Лідиних рук; другу чарку; незвичне ім'я; цілком притомно; хочеться закричати
10	видихнув вільгота; навчити розуму; нашому селі; ваша теорія; попід вікнами	найменшого бажання; головним чином; зачинив двері; найкращим чином; чином пов'язана	мільярди років; студентських років; схиливши голову; підвівши голову; гаком років	центрі міста; ледь чутно; ледь устигла; ледь помітну; вдихнути повітря
9	формі серця; листів од; вільного часу; чомуś винуваті; червоний командир	усім тілом; цілком доречно; цілком природно; обтерся чому; жодному іншому	порветься ніколи; одним боком; цій порі; ту масу; взяти владу	прошепотіла вихована; довга історія; нашої компанії; Нані Новаковської; вигукнула кука
8	могли дозволити; була жахлива; запитала Інна; роби своє; своє гніздечко	адміністратора далі; пропонував усім; цьому металевому; читати вірші; і* тебе	ані крихти; почав кричати; дві тисячі; кому дають; простягнула руку	сто доларів; посеред пустелі; перехрестилася по- думки; намагалася уявити; щосили намагалася
7	Роуз повинна; дайно мені; була втомлена; була встелена; була пауза	усіх сил; руці склянку; ці слова; ці фокуси; ці ніжні	ото бачите; кожного дня; був безсилий; був заставлений; сині очі	можу позбутися; день гніву; першу зустріч; мамки нема; справді гарненька
6	посеред обійстя; мало значення; дитино моя; майнуло щось; летить назустріч	повітря пахне; перестає бути; виграти час; чекало те; хвилини десять	цього поставиться; головою об; моє прізвище; піде геть; йди геть	кінець світу; одним словом; шкода шкода; мій друг; над душею
5	б помилитися; промовив містер; мого двору; через цвинтар; через півгодини	цю здатність; моркви то; мій тато; якусь іншу; нас цікавлять	ані разу; одне слово; між іншими; подумай сама; таке український	Стас здригнувся; зі сну; зі старшого; серцем або; став моїм
4	надто часто; однак	краще ніж; тільки вда-	пішли ми; любили ми;	стільки скільки;

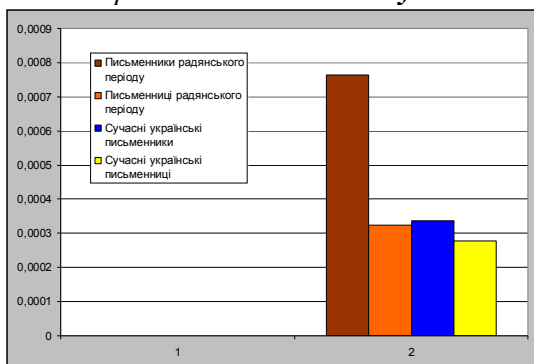
	навколо; горе яке; зостався сам; дуже хотілося	вала; бачив тисячі; просто випадкові; всі пасажери	ми пили; зайшли ми; ми українці	нарешті зміг; двері ногою; сподіваюся ви; вони приймали
3	щось коїться; свої гроші; коло них; де працював; знайдеш де	потрібен цей; твого життя; потягнула мене; Ібрагім був; дім був	вся ця; по пальцях; була рада; від поту; привіт від	зараз навряд; життя справжнє; коридором завжди; б написати; біля Ліди
2	до воріт; дала йому; показуючи йому; скажімо про; було сказано	потрібна ні; дав мені; мав бажання; але стримався; вже двічі	можеш бути; в уяві; тут же; він узяв; він злякався	номер Нані; відкинув із; застиг із; із виглядом; із радянських
1	поруч з; з якимсь; тоді Меланка; особливо тоді; було тихо	це а; слухай ти; потім зробив; ах так; це німецькою	ось де; Данило був; тепер бачу; тепер мусить; чула про	зрозуміла чого; чого власне; в кожному; до краю; думаю що

У подальших дослідженнях заплановано також проаналізувати вербалізації культурнозначущих концептів, оскільки попередній аналіз виявив певні відмінності в уживанні низки слів, зокрема *серце*, *душа*, *тіло*. На діаграмах 14, 15, 16 представлено результати появи слів *серце*, *душа*, *тіло* у складі 2-грам. Здобуті результати засвідчують «часові» і гендерні відмінності у функціонуванні цих слів.

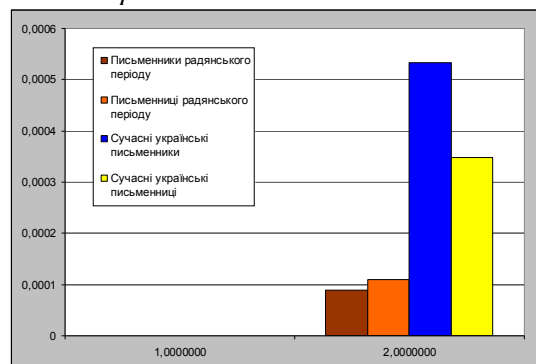
Діаграма 14. Відносна частотність 2-грам із компонентом *серце*



Діаграма 15. Відносна частотність 2-грам із компонентом *душа*



Діаграма 16. Відносна частотність 2-грам із компонентом *тіло*



2-грама «я не» входить до складу низки 3-грам (у Таблиці 8. 3-грами, до яких входить 2-грама я не уміщено дані про праве оточення), серед них у текстах радянських письменників найвищу частотність мають 3-грами я не знаю; я не можу, я не зможу; я не хочу, я не хотів, я не захотіла; у текстах сучасних письменників – я не можу, я не міг; я не знаю, я не знав, я не знала; я не мав, я не маю, я не мала; у текстах радянських письменниць – я не знаю, я не знав, я не знала; я не можу, я не міг; я не хочу, я не хотів, я не хотіла; у текстах сучасних письменниць – я не знаю, я не знав, я не знала; я не можу, я не міг; я не хочу, я не хотіла.

Таблиця 8. 3-грами, до яких входить 2-грама я не

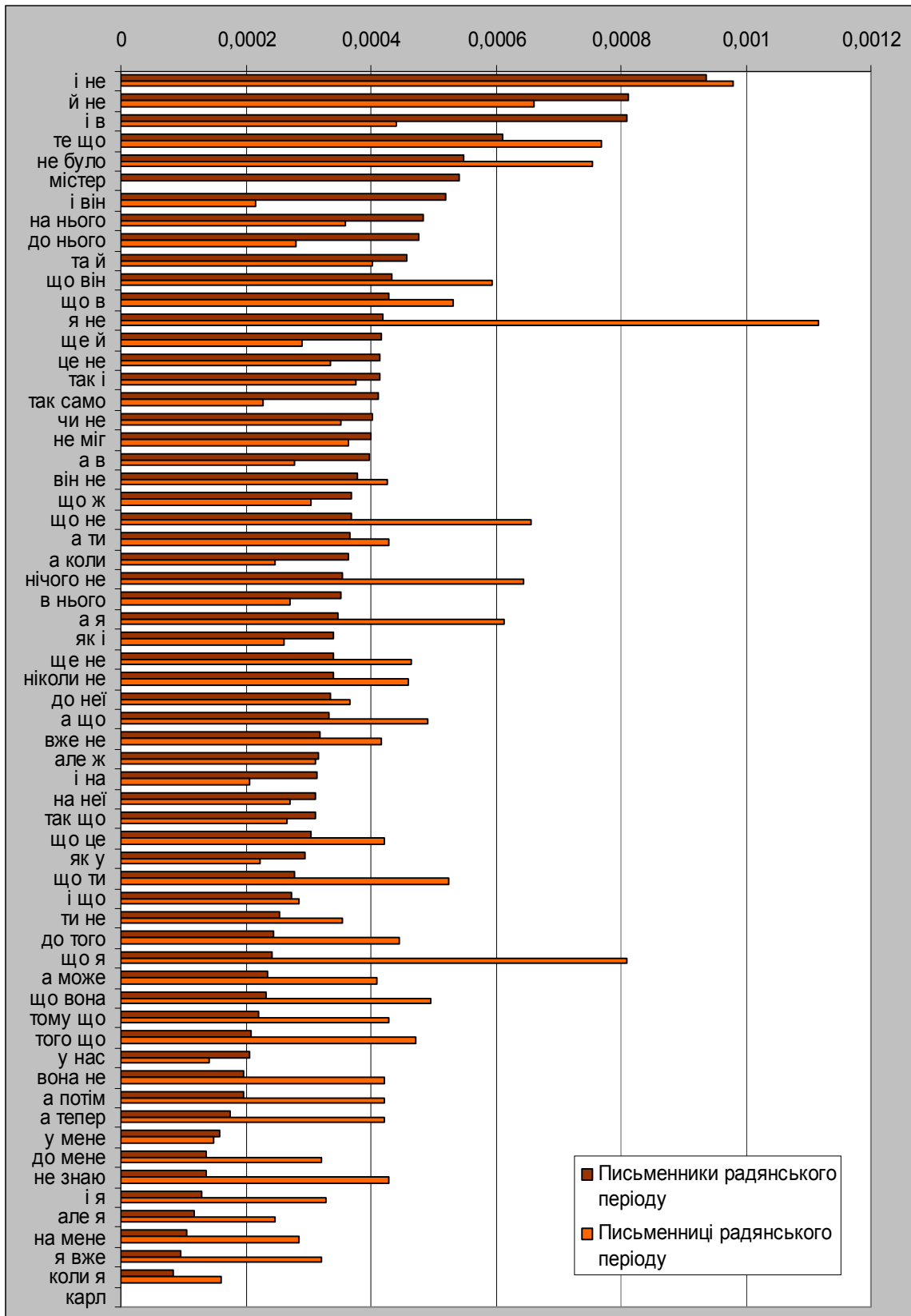
Радянські письменники	Відносна частотність ( $10^{-3}$ )	Радянські письменниці	Відносна частотність ( $10^{-3}$ )	Сучасні укр. письменники	Відносна частотність ( $10^{-3}$ )	Сучасні укр. письменниці	Відносна частотність ( $10^{-3}$ )
я не можу, я не зможу, я не могла	0,035774	я не можу, я не могла, я не міг	0,08593	я не можу, я не могла, я не міг	0,187936	я не можу, я не могла, я не міг	0,0911617
я не знаю, я не знала	0,0310005	я не знаю, я не знав, я не знала	0,104695	я не знаю, я не знав, я не знала	0,128828	я не знаю, я не знав, я не знала	0,080136
я не хочу, я не хотів, я не захотіла	0,0287820	я не хочу, я не хотів, я не хотіла	0,151227	я не хочу, я не хотів, я не хотіла	0,064414	я не хочу, я не хотіла	0,049815
я не з	0,0177120	я не з	0,0078811	–	–	я не з	0,006498
я не сказав, я не кажу	0,0132840	я не скажу, я не кажу	0,0195	я не сказав	0,006441	я не скажу	0,006498
я не маю	0,0110700	я не маю	0,0465310	я не мав, я не маю, я не мала	0,096620	я не маю, я не мала	0,019493
я не збираюся	0,0110700	я не збираюся	0,0078811	я не збираюся, я не збирався	0,012882	я не збираюся	0,006498
я не звик	0,0088560	–	–	–	–	–	–
я не проти	0,0066420	–	–	я не проти	0,006441	я не проти	0,015161
я не чую	0,0044717	–	–	я не чула, я не чув	0,012882	–	–
я не п'ю	0,0044717	–	–	я не п'ю	0,010736	–	–
я не думав	0,0044717	–	–	я не думаю	0,006441	–	–
я не буду	0,0044717	я не буду, я не була	0,0195	я не буду	0,012883	я не була, я не був	0,015161
–	–	–	–	я не став	0,006441	–	–
–	–	я не розумію	0,0232660	я не розуміла	0,006441	я не розумію, я не	0,030322

						зрозуміла, я не зро- зумів	
–	–	–	–	я не раз	0,008588	я не раз	0,006498
–	–	–	–	я не поми- лився	0,008588	я не поми- лилася	0,006498
–	–	–	–	я не певен	0,008588	–	–
–	–	–	–	я не надто	0,008588	–	–
–	–	–	–	я не мушу	0,006441	–	–
–	–	я не люблю	0,00117	я не люблю	0,012883	я не люблю	0,028156
–	–	–	–	я не знайшов	0,008588	–	–
–	–	–	–	я не вмю	0,010736	–	–
–	–	я не вірю	0,00078	я не вірю	0,027913	я не вірю	0,008663
–	–	–	–	я не відразу	0,012883	–	–
–	–	я не витримав	0,00078	я не вит- римала, я не витри- мав	0,017176	я не витримав	0,006498
–	–	–	–	я не бачив	0,015030	я не бачу, я не бачив	0,012996
–	–	я не така	0,00078	–	–	я не така	0,008663
–	–	–	–	–	–	я не дуже	0,008663
–	–	–	–	–	–	я не шиз	0,006498
–	–	–	–	–	–	я не сам	0,006498
–	–	–	–	–	–	я не плакала	0,006498
–	–	–	–	–	–	я не мар- нославний	0,006498

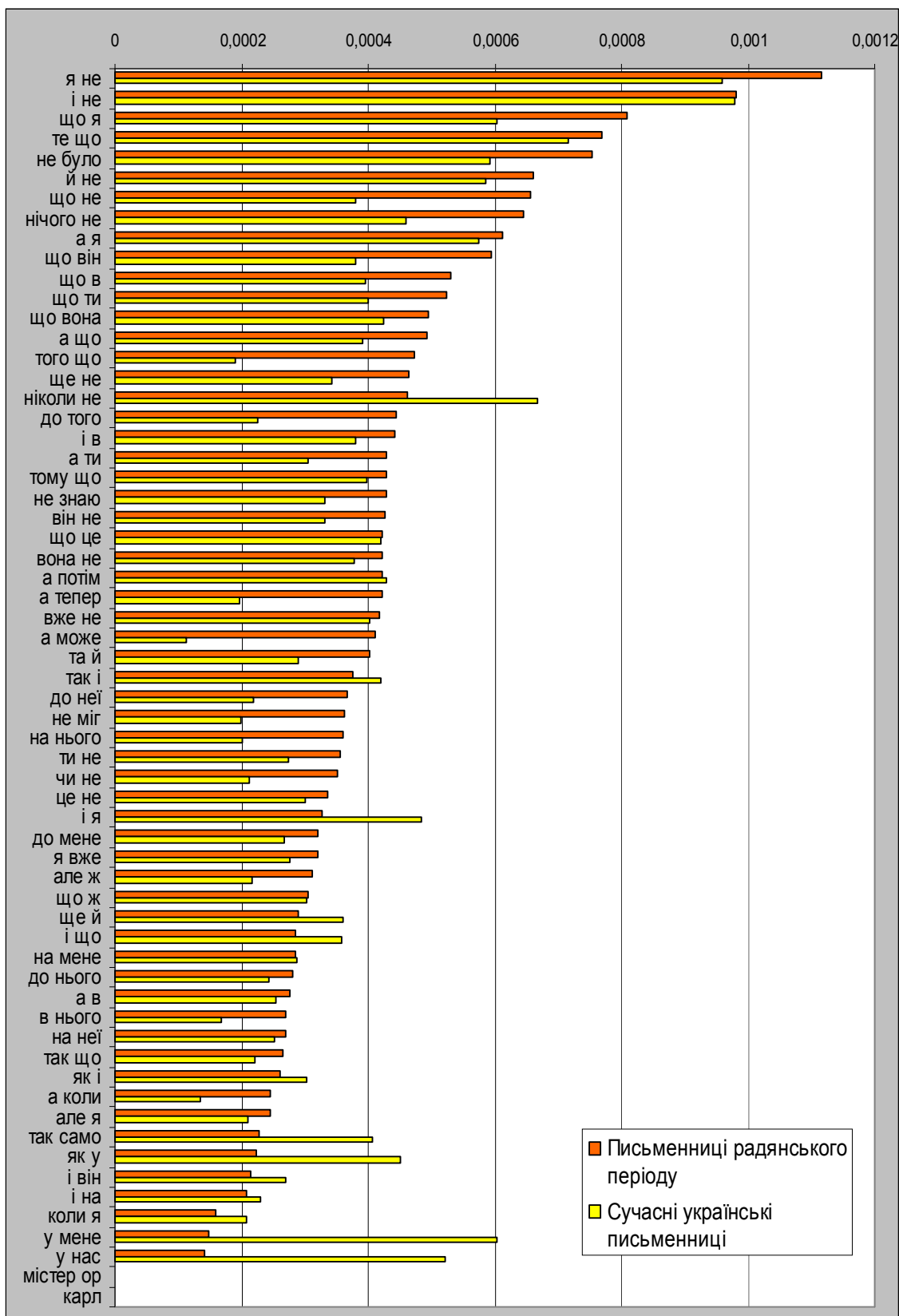
Звертають на себе увагу кілька фактів, зокрема частотність уживання певних сполук письменниками і письменницями радянського періоду (дані подано на Діаграмі 17. *Уживання сполук письменниками і письменницями радянського періоду*). Жінки значно частіше, ніж чоловіки вживають сполуки із запереченням, як-от: *не було, я не, що не, нічого не, ніколи не* тощо. Також частіше в «жіночих» творах радянського періоду трапляються сполуки із займенником *я*.

Дещо інші дані отримано на сучасному матеріалі (див. Діаграма 18. *Зіставлення відносної частотності сполук у творах радянських і сучасних письменниць*): на сполуку *я не* (як і сполуки *що я, і я, тому що, коли я та ін.*) частіше натрапляємо в текстах сучасних українських письменників, ніж письменниць. Натомість у сучасних творах на сполуку *я не* (й інші сполуки із займенником *я*) натрапляємо в «чоловічих» текстах. Жінки частіше послуговуються подвійним запереченням *ніколи не*. Також частіше жіночі тексти містять порівняльну конструкцію *як у*.

Діаграма 17. Уживання сполук письменниками і письменницями радянського періоду



Діаграма 18. Зіставлення відносної частотності сполук у творах радянських і сучасних письменниць





Дослідження n-грам може дати цікаві дані для вивчення, окрім фразеологізмів, жаргонізмів та термінологічних сполук, інтертекстуальності. Колокаціям *як кажуть*, *як то кажуть*, *як казав*, *як казала*, які є маркерами інтертекстом, не властива висока частотність у досліджених текстах, однак аналіз великого масиву текстів дасть змогу дійти висновків про особливості тих чи інших текстів (див. Таблиця 9. *Відносна частотність маркерів інтертекстом*).

Таблиця 9. *Відносна частотність маркерів інтертекстом*

Підкорпуси	Відносна частотність			
	<i>як кажуть</i>	<i>як то кажуть</i>	<i>як казав</i>	<i>як казала</i>
Радянські письменники	0,0000067076	0,0000089	0	0
Радянські письменниці	0,00003515199	0,0000580	0	0
Сучасні укр. письменники	0,00001296106	0,0000043	0	0
Сучасні укр. письменниці	0,0000217051	0	0,0000173641	0,0000173641

Так, у текстах сучасних письменниць трапляються такі інтертекстеми: *А президент скрушно хитає головою, констатує кризу української армії, бідкається, що вона вже стала небезпечною для власного народу, **холера ясна**, як кажуть поляки, то він Головнокомандувач цієї армії чи хто?!; Ви не дослухали, юначе. Знаєте, як кажуть: **той, хто попереджений озброєний!**; Або він бабій, який, як кажуть, **ціну собі не складе**, або педант, від якого жінки тікають, мов чорт від ладану, або: ну, в нього не все в порядку по чоловічій частині; І, як кажуть японці, **знайти своє обличчя**; Йому це, як кажуть, **по барабану**; Жінка навпаки кине весь цей мотлох заради того, щоб, як кажуть, “**знайти рай у курені**”; ... Стодоля ніколи не бовкав зайвого й не дався ідентифікувати), чувся за тим радше жаль, що і в такого великого чоловіка, як полковник Коновалець, **знайшлася слабкість**, хай навіть така **дрібонька, єк синє за нігтем**, як кажуть тутешні селяни... та ін.*

Також сформовано переліки 30 найчастотніших 3-грам у кожному з підкорпусів (у Таблиці 10. *Найчастотніші 3-грами* подано здобуті результати, риска означає, що така 3-грама не ввійшла до списку 30 найчастотніших). До найчастотніших триграм також потрапили оніми – імена героїв художніх творів (*Карл Йозеф Цумбруннен, Іветта Андріївна Вербицька*).

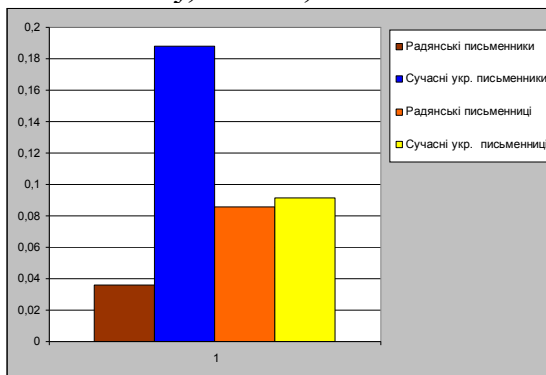
Таблиця 10. Найчастотніші 3-грами

3-грами	Відносна частотність			
	Радянські письменниці	Радянські письменники	Сучасні укр. письменниці	Сучасні укр. письменники
до того ж	0,000081430	0,000084132	0,000071473	0,000133121
про те що	0,000089185	0,000110700	0,000160273	0,000124533
так і не	0,000089185	0,000077490	0,000138615	0,000122386
в тому що	0,000100818	–	0,000047649	0,000083738
я не знаю	0,000065919	–	0,000056312	0,000081591
про те як	–	–	0,000056312	0,000081591
я не можу	0,000069797	–	0,000054146	0,000081591
я не міг	–	–	–	0,000077296
що я не	0,000093063	–	–	0,000073002
Карл Йозеф Цумбруннен	–	–	–	0,000073002
те що я	–	–	0,000062810	0,000070855
час від часу	–	–	0,000054146	0,000070855
після того як	–	0,000048708	0,000062810	0,000066561
я подумав що	–	–	–	0,000064414
але я не	0,000069797	–	0,000047649	0,000062267
я вже не	0,000062042	–	0,000043317	0,000062267
дивиться на мене	–	–	–	0,000062267
так само як	–	0,000101844	0,000071473	0,000060119
я не мав	–	–	–	0,000060119
для того щоб	0,000089185	0,000092988	0,000043317	0,000055825
і я не	–	–	0,000043317	0,000053678
перед тим як	–	–	0,000051981	0,000051531
на те що	–	0,000037638	–	0,000051531
за те що	0,000093063	0,000075276	0,000060644	0,000049384
а тим часом	–	–	–	0,000049384
до мене і	–	–	–	0,000049384
і коли я	–	–	–	0,000049384
для того аби	–	–	–	0,000047237
не знав що	–	–	–	0,000047237
таке враження що	–	–	–	0,000047237
та ще й	–	0,000042066	0,000071473	–
я ніколи не	–	–	0,000056312	–
те що вона	–	–	0,000054146	–
раз у раз	–	0,000061992	0,000051981	–
за десять хвилин	–	–	0,000047649	–
Іветта Андріївна Вербицька	–	–	0,000047649	–
у мене є	–	–	0,000047649	–
а що як	–	–	0,000045483	–
все ж таки	0,000182248	0,000046494	0,000043317	–
я знаю що	0,000062042	–	0,000043317	–
я більше не	–	–	0,000043317	–
про це не	–	–	0,000041151	–
не те що	0,000096940	–	0,000038985	–
в цей час	–	0,000099630	–	–
а може й	–	0,000055350	–	–
і вже не	–	0,000050922	–	–
та все ж	–	0,000050922	–	–
містер Ор не	–	0,000048708	–	–
що ж це	–	0,000046494	–	–
можна було б	–	0,000044280	–	–
ніяк не міг	–	0,000044280	–	–

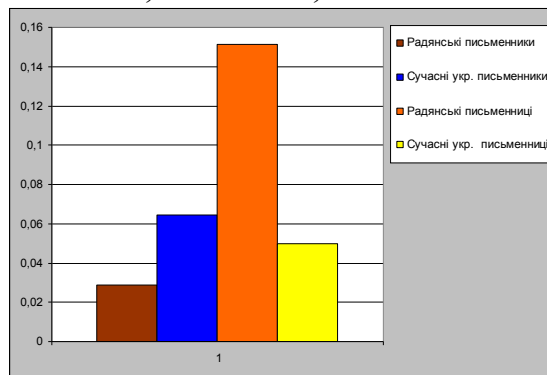
а де ж	–	0,000042066	–	–
а що ж	–	0,000042066	–	–
вже й не	–	0,000042066	–	–
ми з тобою	–	0,000042066	–	–
до того що	0,000093063	0,000039852	–	–
а чого ж	–	0,000037638	–	–
й не було	–	0,000037638	–	–
та й не	–	0,000037638	–	–
тільки тому що	–	0,000037638	–	–
ще й досі	–	0,000037638	–	–
що він не	–	0,000037638	–	–
від того що	0,000112451	–	–	–
я не хочу	0,000112451	–	–	–
з того що	0,000108573	–	–	–
мені здається що	0,000093063	–	–	–
не може бути	0,000073675	–	–	–
а все ж	0,000065919	–	–	–
перш за все	0,000065919	–	–	–
а я не	0,000062042	–	–	–
прошу я вас	0,000062042	–	–	–
в той час	0,000058164	–	–	–
не знаю що	0,000058164	–	–	–
не можна було	0,000058164	–	–	–
що з того	0,000058164	–	–	–
як то кажуть	0,000058164	–	–	–
я ще не	0,000054287	–	–	–

Аналіз 3-грам засвідчує той факт, що найчастотнішими 3-грамами, які наявні в кожному з досліджуваних підкорпусів, є: *про те що, для того щоб, до того ж, так і не, за те що*. Симптоматичними є результати появи в досліджуваних текстах 3-грам *я не знаю (я не знав, я не знала); я не можу (я не зможу, я не міг, я не могла); я не хочу (я не хотів, я не хотіла, я не захотіла)*. Так, наприклад, на Діаграмах 19, 20, 21 показано зіставлення уживання 3-грам *я не можу (я не зможу, я не міг, я не могла); я не хочу (я не хотів, я не хотіла, я не захотіла)*.

Діаграма 19. 3-грами *я не можу, я не зможу, я не міг, я не могла*



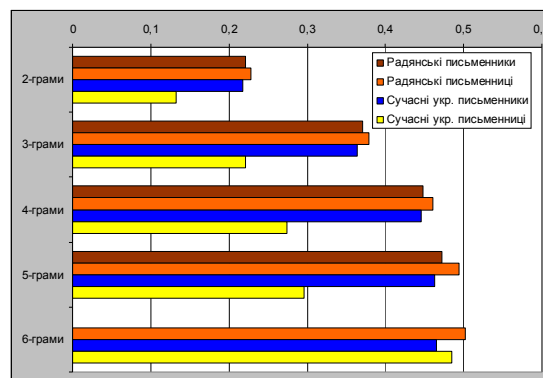
Діаграма 20. 3-грама *я не хочу, я не хотів, я не хотіла, я не захотіла*



Найчастіше вислів *я не можу* трапляється в текстах сучасних українських письменників та письменниць. Про своє небажання (вислів *я не хочу*) найчастіше заявляють радянські письменниці, а найрідше згаданий вислів ужито у творах радянських письменників.

Окрім того, визначено показник різноманітності *n*-грам, дані про який представлено на Діаграмі 21. Показник різноманітності *n*-грам у досліджуваних підкорпусах. Найвищий показник різноманітності 2-грам властивий текстам радянських письменниць, а найнижчий – текстам сучасних українських авторок.

Діаграма 21. Показник різноманітності *n*-грам у досліджуваних підкорпусах



Т. Бобкова дійшла висновку, що загальний відсоток колокацій, уживаних принаймні 50 разів, у законодавчих текстах перевищує 2%, натомість відповідні за частотою стійкі сполуки публіцистичних і наукових текстів становлять менше 1% (0,96% і 0,77%) [Бобкова 2015: 40]. Згадана авторка виявила, що в публіцистичних текстах найчастотнішою є колокація *в, Україна* (1249), а в наукових – *український, мова* (554), в законодавчих – *відповідно до* (2361) [Бобкова 2015: 40]. Дані нашого дослідницького корпусу засвідчують той факт, що колокація *в Україні* трапляється 10 разів (0,00215%), *в Україну* – (0,00193%), *потрібно в Україну* – 3 (0,0006441%) у текстах сучасних українських письменників; *в Україні* – 13 (0,0028156); *в Україну* – 5 (0,0010829%) у текстах сучасних українських письменниць. Натомість в текстах радянських письменників двічі виявлено сполуку *на Україні* (0,0000044717), 1 раз – *на Україну*. Колокація *українською мовою* трапляється лише 3 рази (0,00064) у текстах сучасних українських письменників. На словосполуку *відповідно до* натрапляємо 1 раз у текстах сучасних українських письменників, 2 рази – в текстах сучасних українських письменниць (0,0000043410). Зіставлення результатів Т. Бобкової з результатами,

отриманими на матеріалі текстів художнього стилю, підтверджує припущення про те, що колокації можуть слугувати як маркерами функціональних стилів, так і маркерами ідостилів.

Отже, застосування алгоритму «Суфіксне дерево» розподілу тексту на кластери, або так звані n-грами, дало змогу розподілити досліджуваний масив текстів на 2-6-грами. Більша різноманітність лексики властива загалом сучасним авторам й авторкам порівняно з письменниками й письменницями радянського періоду, при чому сучасним «жіночим» текстам притаманна більша різноманітність лексики, ніж сучасним «чоловічим».

Аналіз n-грам виявив, що найвищу відносну частотність у всіх досліджених підкорпусах мають 2-грами (їхня відносна частотність коливається в діапазоні 0,351-0,388). Зі збільшенням кількості компонентів n-грами відносна частотність зменшується. Найбільше 2-грам мають відносну частотність у діапазоні 0,0001-0,0009. Загалом сучасні художні тексти, зокрема чоловічі, містять більше колокацій, ніж тексти радянського періоду.

До найчастотніших 2-грам належать сполуки формальної будови *сполучник + заперечна частка*, *сполучник + прийменник*, *прийменник + займенник*, *прийменник + заперечна частка*, *сполучник + сполучник*, *дієслово + заперечна частка*, *дієслово + прийменник* тощо. Висока відносна частотність властива 2-грамам *я не* (найвища відносна частотність характерна для текстів сучасних письменників, натомість 2-грама *не знаю* найчастіше трапляється в текстах радянських письменниць).

Підрахунки коефіцієнту VI підтверджують те, що найчастотнішим 2-грамам властивий низький коефіцієнт VI. У діапазоні коефіцієнта VI від 1 до 29 найчастотнішими є 2-грами з коефіцієнтом VI, який дорівнює 1.

У досліджених підкорпусах 2-грами з найвищим коефіцієнтом VI здебільшого мають формальну структуру *прикметник + іменник*, *іменник + іменник*. Як уже зазначалося, докладно проаналізувати формальну структуру заплановано в подальшому.

У досліджених текстах суттєві відмінності спостережено в уживанні чоловіками та жінками таких слів, як *серце* (висока відносна частотність 2-грам із цим компонентом властива жіночим текстам різних періодів); *душа* (висока відносна частотність характерна для текстів радянських письменників); *тіло* (частотне уживання притаманне для сучасних українських письменників).

Певних висновків про світосприймання авторів та авторок різних часових періодів можна дійти на основі аналізу таких 2-грам, як: *не було, я не, що не, ніколи не*, які властиві жіночим текстам радянського періоду, як порівняти з «чоловічими» текстами цього періоду. Натомість висока частотність сполуки *я не* характерна для сучасних чоловічих текстів. Аналогічні висновки стосуються 3-грам: сполуки *я не можу, я не зможу, я не міг, я не могла* найчастіше використовують сучасні українські письменники; сполуки *я не хочу, я не хотів, я не хотіла, я не захотіла* – радянські жінки.

Вища різноманітність n-грам притаманна радянським авторам та авторкам.

У подальшому належить збільшити обсяг досліджуваного матеріалу, використати інші методи статистичного аналізу n-грам для верифікації здобутих результатів, які вважаємо попередніми.

#### Література

Алексеенко 1999: Алексеенко, М. А.: Негативные стереотипы базового концепта культуры «женщина» в паремиологическом выражении. In: Русский язык, литература и культура на рубеже веков, Братислава 1999, Ч. 1, с. 13–14.

Алексеенко 2000: Алексеенко, М.: Некоторые аспекты вербализации концепта культуры «женщина». In: Slavica Stetinensia. №10. 2000, с. 109–122.

Алексеенко 2005: Алексеенко, М.: О асимметрии полового диморфизма в современном русском языке. In: Slowo. Tekst. Czas, Szczecin 2005, с. 205–213.

Алексеенко 2006: Алексеенко, М. А.: Негативен ли образ женщины в языке? In: Слово в словаре и дискурсе, Москва 2006, с. 22-30.

Арсентьева та ін. 2009: Арсентьева, Е. Ф., Коноплёва, Н. В.: Семантические особенности фразеологизмов с маскулинной гендерной референцией в английском и русском языках In: Вестник КГУ им. Н.А. Некрасова. №2. 2009. [Електронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/semanticheskie-osobennosti-frazeologizmov-s-maskulinnoy-gendernoy-referentsiyev-angliyskom-i-russkom-yazykah>, 29.10.2015.

Архангельская 2005: Архангельская, А.: Маскуліноцентричність мови як образ світу, у слові явлений. In: Slowo. Tekst. Czas, Szczecin 2005, s. 326–330.

Архангельская 2011: Архангельская, А.: Сексизм в языке: мифы и реальность. Оломоуц 2011.

Барвіна 2013: Барвіна, Н. О.: Моделювання концепту на діалектному матеріалі: Концепт Жінка в українській фразеології: культурні стереотипи. In: Лінгвістика. №2(29). 2013, с. 108-114.

Белая 2009: Белая, Н. В.: О проявлении гендерных различий в языке In: Вестник МГТУ им. М. А. Шолохова. Филологические науки. №3. 2009. [Електронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/o-proyavlenii-gendernyh-razlichiy-v-yazyke>, 29.10.2015.

Бобкова 2015: Бобкова, Т.: Статистичні ознаки колокацій як маркери функціональних стилів. In: Українське мовознавство. Вип. 41/1. 2015, с. 38-43.

Васильева 2013: Васильева, И. Б.: Смягчающие операторы степени английского языка: гендерный аспект In: Вестник ЧелГУ. №20 (311). 2013. [Електронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/smyagchayuschie-operatory-stepeni-angliyskogo-yazyka-gendernyy-aspekt>, 29.10.2015.

Викторова 2011: Викторова, Е. Ю.: Влияет ли гендер на использование дискурсивов? (На материале письменного научного дискурса) In: Изв. Сарат. ун-та Нов. сер.

Сер. Филология. Журналистика. 2011. №3. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/vliyaet-li-gender-na-ispolzovanie-diskursivov-na-materiale-pismennogo-nauchnogo-diskursa>, 28.10.2015.

Гацалова та ін. 2010: Гацалова, Л. Б., Парсиева, Л. К.: Этноспецифика эмотивных устойчивых сочетаний в осетинском языке In: Вестник ЧелГУ. №29. 2010. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/etnospetsifika-emosivnyh-ustoychivyh-sochetaniy-v-osetinskom-yazyke>, 29.10.2015.

Гибадуллина 2010: Гибадуллина, Л. З.: Языковые особенности современной женской прозы (гендерный аспект) In: Вестник Башкирск. ун-та. №3-1. 2010. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/yazykovye-osobennosti-sovremennoy-zhenskoj-prozy-gendernyy-aspekt>, 29.10.2015.

Голубева 2009: Голубева, С. С.: Лингвистическая маркированность гендера как социокультурной категории In: Вестник ЮУрГУ. Серия: Лингвистика. №25(158). 2009. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/lingvisticheskaya-markirovannost-gendera-kak-sotsiokulturnoy-kategorii>, 28.10.2015.

Гончар 2009: Гончар, Ю. О.: Гендерна інтерпретація художнього світу Тараса Шевченка: автореф. дис... канд. філол. наук: 10.01.01. Кіровоград. держ. пед. ун-т ім. В.Винниченка. Кіровоград 2009.

Горина 2011: Горина, О. Г.: Применение методов корпусной лингвистики для определения контекстно-специфических слов и коллокаций. In: Вестник ЛГУ им. А. С. Пушкина. №3. 2011. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/primeneniye-metodov-korpusnoy-lingvistiki-dlya-opredeleniya-kontekstno-spetsificheskikh-slov-i-kollokatsiy>, 04.01.2016.

Горошко 2009: Горошко, Е.И.: Информационно-коммуникативное общество в гендерном измерении: Монография. Харьков 2009.

Горошко 2010: Горошко, О. I.: Гендерні аспекти інтернет-комунікацій: автореф. дис. ... д-ра соціол. наук : 22.00.04. Харк. нац. ун-т ім. В.Н. Каразіна. Харків 2010.

Горошко: Горошко, Е. И.: Гендерные исследования в языкознании (к проблеме становления метода) In: Культура народов Причерноморья, Межвузовский центр Крыма. №45. т.1. 2004. [Электронна версія]. Режим доступу: <http://textology.ru/article.aspx?aId=83>, 28.10.2015.

Гриценко та ін. 2012: Гриценко, Е. С., Сергеева, М. В., Барина, А. О.: Динамика конструирования этничности и гендера в англоязычном лексикографическом дискурсе. In: Вестник МГЛУ. №638. 2012. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/dinamika-konstruirovaniya-etnichnosti-i-gendera-v-angloyazychnom-leksikograficheskom-diskurse>, 28.10.2015.

Демидова 2011: Демидова, Т. Д.: Периферийная часть лексикона как показатель литературной манеры писателя (на материале лирических миниатюр В. П. Деткова). In: Вестник ЛГУ им. А. С. Пушкина. №2. 2011. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/periferiyayaya-chast-leksikona-kak-pokazatel-literaturnoy-manery-pisatelya-na-materiale-liricheskikh-miniatur-v-p-detkova>, 03.01.2016.

Елемешина 2010: Елемешина, Ю. А.: Проявление гендерных различий на сегментном и суперсегментном уровнях языка. In: Вестник МГЛУ. №592. 2010. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/proyavlenie-gendernyh-razlichiy-na-segmentnom-i-suprasegmentnom-urovnyah-yazyka>, 29.10.2015.

Ермакова 2014: Ермакова, Л. А.: Изучение поэтического языка М. Цветаевой в гендерном аспекте: проблемы и перспективы. In: Вестник КГУ им. Н. А. Некрасова. №5. 2014. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/izuchenie-poeticheskogo-yazyka-m-tsvetaevoj-v-gendernom-aspekte-problemy-i-perspektivy>, 29.10.2015.

Ермолаева 2008: Ермолаева, М. В.: Фразеологические новообразования, реализующие поэтические установки гендера. In: Вестник ЧГАКИ. №3(15). 2008. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/frazeologicheskie-novoobrazovaniya-realizuyushchie-pragmaticheskie-ustanovki-gendera>, 28.10.2015.

Жеребкин 2006: Жеребкин, С.: Наслаждение быть украинцем: украинская маскулинность и Оранжевая революция. In: Гендерные исследования. №14. 2006, с. 226-245.

Захаров 2015: Захаров, В.П.: Сочетаемость через призму корпусов. In: Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 27 — 30 мая 2015 г.). Вып. 14 (21). Т. 1. Москва 2015. С. 667-682. [Электронна версія]. Режим доступу: [https://www.academia.edu/15102853/%D0%A1%D0%BE%D1%87%D0%B5%D1%82%D0%B0%D0%B5%D0%BC%D0%BE%D1%81%D1%82%D1%8C\\_%D1%87%D0%B5%D1%80%D0%B5%D0%B7\\_%D0%BF%D1%80%D0%B8%D0%B7%D0%BC%D1%83\\_%D0%BA%D0%BE%D1%80%D0%BF%D1%83%D1%81%D0%BE%D0%B2\\_Set\\_Phrases\\_a\\_View\\_through\\_Corpora](https://www.academia.edu/15102853/%D0%A1%D0%BE%D1%87%D0%B5%D1%82%D0%B0%D0%B5%D0%BC%D0%BE%D1%81%D1%82%D1%8C_%D1%87%D0%B5%D1%80%D0%B5%D0%B7_%D0%BF%D1%80%D0%B8%D0%B7%D0%BC%D1%83_%D0%BA%D0%BE%D1%80%D0%BF%D1%83%D1%81%D0%BE%D0%B2_Set_Phrases_a_View_through_Corpora), 28.10.2015.

Захаров, Хохлова 2014: Захаров, В.П., Хохлова, М.В.: Автоматическое выявление терминологических словосочетаний. In: Структурная и прикладная лингвистика. Вып.10. Санкт-Петербург 2014. С.182-200. [Электронна версія]. Режим доступу: [https://www.academia.edu/8587168/%D0%90%D0%92%D0%A2%D0%9E%D0%9C%D0%90%D0%A2%D0%98%D0%A7%D0%95%D0%A1%D0%9A%D0%9E%D0%95\\_%D0%92%D0%AB%D0%AF%D0%92%D0%9B%D0%95%D0%9D%D0%98%D0%95\\_%D0%A2%D0%95%D0%A0%D0%9C%D0%98%D0%9D%D0%9E%D0%9B%D0%9E%D0%93%D0%98%D0%A7%D0%95%D0%A1%D0%9A%D0%98%D0%A5\\_%D0%A1%D0%9B%D0%9E%D0%92%D0%9E%D0%A1%D0%9E%D0%A7%D0%95%D0%A2%D0%90%D0%9D%D0%98%D0%99](https://www.academia.edu/8587168/%D0%90%D0%92%D0%A2%D0%9E%D0%9C%D0%90%D0%A2%D0%98%D0%A7%D0%95%D0%A1%D0%9A%D0%9E%D0%95_%D0%92%D0%AB%D0%AF%D0%92%D0%9B%D0%95%D0%9D%D0%98%D0%95_%D0%A2%D0%95%D0%A0%D0%9C%D0%98%D0%9D%D0%9E%D0%9B%D0%9E%D0%93%D0%98%D0%A7%D0%95%D0%A1%D0%9A%D0%98%D0%A5_%D0%A1%D0%9B%D0%9E%D0%92%D0%9E%D0%A1%D0%9E%D0%A7%D0%95%D0%A2%D0%90%D0%9D%D0%98%D0%99), 28.10.2015.

Касумова та ін. 2013: Касумова, М. Ю., Мерве Сердар: Специфика гендерной коммуникации в Интернет-среде. In: Актуальные проблемы гуманитарных и естественных наук. №5. 2013. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/spetsifika-gendernoy-kommunikatsii-v-internet-srede>, 28.10.2015.

Кирилина 1999: Кирилина, А. В.: Гендер: лингвистические аспекты, Москва 1999.

Кирилина 2004: Кирилина, А. В.: Гендерные исследования в лингвистике и теории коммуникации: уч. пособие. Москва 2004.

Комов 2012: Комов, О. В.: Морфологічні особливості мовлення чоловіків і жінок (комп'ютерний дискурс). In: Дослідження з лексикології і граматики української мови. Вип. 12. 2012. [Электронна версія]. Режим доступу: [http://nbuv.gov.ua/UJRN/dlgum\\_2012\\_12\\_5](http://nbuv.gov.ua/UJRN/dlgum_2012_12_5), 28.10.2015.

Компанцева 2006: Компанцева, Л. Ф.: Гендерные основы Интернет-коммуникации в постсоветском пространстве: Моногр. Луган. нац. пед. ун-т им. Т. Шевченко. Альма-матер, Луганск 2006.

Корольова 2014: Корольова, В. В.: Гендерні концепти в мовній свідомості сучасної молоді. In: Дослідження з лексикології і граматики української мови. Вип. 15. 2014. С. 11-17. [Электронна версія]. Режим доступу: [http://nbuv.gov.ua/UJRN/dlgum\\_2014\\_15\\_4](http://nbuv.gov.ua/UJRN/dlgum_2014_15_4), 28.10.2015.

Кочеткова 2013: Кочеткова, Н. А.: Статистические языковые методы. Коллокации и коллигации. In: Новые информационные технологии в автоматизированных системах. №16. 2013. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/statisticheskie-yazykovye-metody-kollokatsii-i-kolligatsii>, 10.01.2016.

Левченко 2008: Левченко, О.: Стереотип чоловіка в жіночих прецедентних текстах та фразеології. In: Słowo. Tekst. Czas., Szczecin 2008, s. 61–68.

Левченко 2010: Левченко, Е.: Гендерный аспект компаративных фразеологических единиц. In: Мовна картина світу словян і культура. Т. 1. Матеріали міжнародної наукової конференції (Люблін, 5-7 листопада 2010 р.) Люблін-Рівне 2010, с. 44-49.

Луцьянова 2009: Луцьянова, Т. Г.: Вільні атрибутивні словосполучення як засіб об'єктивації гендерних стереотипів (на матеріалі сучасного британського газетного дискурсу) : автореф. дис ... канд. філол. наук: 10.02.04. Харків 2009.

Маслова 2011: Маслова, Ю. П.: Гендерний дискурс сучасних друкованих україномовних ЗМІ: автореф. дис. ... канд. філол. наук : 10.02.01. Волин. нац. ун-т ім. Л. Українки. Луцьк 2011.

Мінаєва 2007: Мінаєва, Е. В. Гендерна концептологія: мовна репрезентація концептів "Дім" і "Любов" у жіночій поезії: автореф. дис... канд. філол. наук: 10.02.02. Тавр. нац. ун-т ім. В.І.Вернадського. Сімферополь 2007.

Мухин 2010: Мухин, М. Ю.: Концептуальные профили произведений М. Булгакова, В. Набокова, А. Платонова и М. Шолохова (по данным сопоставительного анализа частотной



лексики. In: Вестник БФУ им. И. Канта. №8. 2010. [Електронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/kontseptualnye-profil-proizvedeniy-m-bulgakova-v-nabokova-a-platonova-i-m-sholohova-po-dannym-sopostavitelnogo-analiza-chastotnoy>, 24.10.2015.

Нгуен Суан Тхе 2011: Нгуен Суан Тхе: Гендерные особенности оценочной семантики во вторичной номинации. In: Известия ТулГУ. Гуманитарные науки. №2. 2011. [Електронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/gendernye-osobennosti-otsenочноy-semantiki-vo-vtorichnoy-nominatsii>, 28.10.2015.

Павлов та ін. 2011: Павлов, Ю. Н., Тихомирова, Е. А.: Оценка устойчивости во времени частотных словарей авторов в задачах идентификации текстов. In: Наука и образование: электронное научно-техническое издание. №12. 2011. [Електронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/77-30569-274006-otsenka-ustoychivosti-vo-vremeni-chastotnyh-slovaray-avtorov-v-zadachah-identifikatsii-tekstov>, 24.10.2015.

Пирматова 2013: Пирматова, М. М.: Экспрессивность и гендер. In: Актуальные проблемы гуманитарных и естественных наук. №8. 2013. [Електронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/ekspressivnost-i-gender>, 28.10.2015.

Пузиренко 2005: Пузиренко, Я. В.: Агентивно-професійні назви осіб жіночої статі в лексикографічному описі та узусі 2005 года: Автореф. дис... канд. філол. наук: 10.02.15. Київ. нац. ун-т ім. Т.Шевченка. Ін-т філол. Київ 2005.

Рудюк 2010: Рудюк, Т. В.: Вербалізація концептів ЧОЛОВІК, ЖІНКА в українській фразеології: автореф. дис. ... канд. філол. наук : 10.02.01. Нац. пед. ун-т ім. М. П. Драгоманова. Київ 2010.

Самохина 2011: Самохина, А. А.: Сравнение частотности употребления стилистически маркированных и стилистически нейтральных терминов в художественном тексте. In: Известия РГПУ им. А. И. Герцена. №131. 2011. [Електронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/sravnenie-chastotnosti-upotrebleniya-stilisticheski-markirovannyh-i-stilisticheski-neytralnyh-terminov-v-hudozhestvennom-tekste>, 24.10.2015.

Ставицька 2003: Ставицька, Л. О.: Актуальні проблеми сучасної української гендерної лінгвістики. [Електронна версія]. Режим доступу: [http://www.linguistics.kiev.ua/seminar/2003/12/18/sem01\\_01\\_6.html](http://www.linguistics.kiev.ua/seminar/2003/12/18/sem01_01_6.html), 28.10.2015.

Ставицька 2005: Ставицька, Л.: Гендерні стереотипи в сучасній мовній свідомості: за даними асоціативного експерименту зі словами мати, батько. In: Дивослово. № 6. 2005, с. 45-50.

Ставицька 2006: Ставицька, Л. О.: «Чоловік (мужчина)» у концептосфері української фразеології. In: Мовознавство 2006, № 2–3, с. 118–129.

Стяжкіна 2015: Стяжкіна, О.: Особистісний світ жінки-літератора в Україні у другій половині ХХ століття: спроба антропологічного аналізу. [Електронна версія]. Режим доступу: <http://uamoderna.com/md/stiazhkina-women-writers-ukraine>, 28.10.2015.

Сукаленко 2007а: Сукаленко, Т. М.: Метафорична парадигма „птаха-жінка” (семасіологічний аспект аналізу гендерного концепту). In: Лінгвістичні дослідження: зб. наук. праць. Вип. 23. Харків 2007, с. 80-87.

Сукаленко 2007б: Сукаленко, Т. М.: До проблеми метафоричного вираження концепту «жінка» в українській мові. In: Лінгвістика: зб. наук. праць. № 1(11). Луганськ 2007, с. 74-81.

Сукаленко 2007в: Сукаленко, Т. М.: Жінка в дзеркалі метафори. In: Вісник Донецького університету. Серія „Б”. Гуманітарні науки. № 1–2. Донецьк 2007, с. 63-70.

Сукаленко 2007г: Сукаленко, Т. М.: Парадигма „жінка – предметний світ” як засіб реалізації концепту „жінка” в українській мові. In: Вісник Черкаського університету. Серія „Філологічні науки”. Вип. 102. Черкаси 2007, с. 118-124.

Сукаленко 2008: Сукаленко, Т. М.: Метафоричний образ жінки у мовній свідомості українців In: Науковий вісник Волинського національного університету імені Лесі Українки. Серія „Філологічні науки”. № 2. Луцьк 2008, с. 54-58.

Сукаленко 2010: Сукаленко, Т. М.: Типізовані образні парадигми словесного втілення концепту жінка в українській мові [Текст]. In: Українська мова : Науково-теоретичний журнал. №3. 2010, с. 81-98.

Сусов 2006: Сусов, И.П.: История языкознания: Учебное пособие для студентов старших курсов и аспирантов. Москва 2006.

Тамерьян 2014: Тамерьян, Т. Ю. Гендерный аспект английской метафорической картины мира. In: Вестник СамГУ. №1 (112). 2014. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/gendernyy-aspekt-angliyskoy-metaforicheskoy-kartiny-mira>, 28.10.2015.

Тараненко 2005: Тараненко, О. О. : Принцип андроцентризму в системі мовних координат і сучасний гендерний рух. In: Мовознавство. № 1. 2005, с. 3-25.

Хачмафова 2010: Хачмафова, З. Р.: Гендерная стратификация языка женской прозы (на материале русского и немецкого языков). In: Известия РГПУ им. А.И. Герцена. №120. 2010. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/gendernaya-stratifikatsiya-yazyka-zhenskoj-prozy-na-materiale-russkogo-i-nemetskogo-yazykov>, 29.10.2015.

Шаманская 2012: Шаманская, М. А.: Лексическая сочетаемость как средство репрезентации гендера в немецком языке. In: Magister Dixit. №4. 2012. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/leksicheskaya-sochetaemost-kak-sredstvo-reprezentatsii-gendera-v-nemetskom-yazyke>, 28.10.2015.

Шишигина 2012: Шишигина, О. Ю.: Репрезентация концепта «MAN» атрибутивными прилагательными в современном варианте английского языка. In: Вестник ЧелГУ. №2. 2012. [Электронна версія]. Режим доступу: <http://cyberleninka.ru/article/n/reprezentatsiya-kontsepta-man-atributivnymi-prilagatelnyimi-v-sovremennom-amerikanskom-variante-angliyskogo-yazyka>, 09.01.2016.

Эмирова 2002: Эмирова, А. М.: Гендер в зеркале русской фразеологии. In: Культура народов Причерноморья. №31. 2002, с. 210–212.

Ягунова 2010: Ягунова, Е.В. Пивоварова, Л.М.: Природа коллокаций в русском языке. Опыт автоматического извлечения и классификации на материале новостных текстов. In: Сб. НТИ. Сер.2. №5. 2010. [Электронна версія]. Режим доступу: [http://jf.spbu.ru/upload/files/file\\_1357769727\\_3479.pdf](http://jf.spbu.ru/upload/files/file_1357769727_3479.pdf), 10.01.2016.

Кuczynska та ін. 2002: Kuczynska, M. , Aleksiejenko, M. : Jejimię znaczy – życie. In: Słowo. Tekst. Czas, Szczecin 2002, s. 133–146.

#### Анотація

У цій праці на основі створеного дослідницького корпусу здійснено спробу виявити колокації, наявні в українській художній прозі різних часових періодів та дослідити особливості вживання колокацій у художніх текстах чоловіків та жінок. Для проведення аналізу було створено дослідницькі підкорпуси: підкорпус художніх текстів 1961-1990 рр., який містить тексти (обсяг становить приблизно 710 тис. слововживань) письменниць та письменників; підкорпус текстів 2000-х рр. (обсяг приблизно 927 тис. слововживань): письменниць та письменників. На матеріалі дослідницьких корпусів здійснено спробу опрацювання методики аналізу колокацій (n-грам). Для аналізу застосовано алгоритм "Суфіксне дерево" розподілу тексту на кластери, або так звані n-грами. У цьому дослідженні загальний масив текстів було розподілено на 2, 3, 4, 5, 6-грами. Здійснено кількісний аналіз. Для виявлення колокацій також застосовано коефіцієнт взаємної інформації. Зроблено висновки про висновки про особливості вживання колокацій у «чоловічих» та «жіночих» текстах різних часових періодів.

**Ключові слова:** колокація, n-грама, корпус текстів, художня проза, алгоритм «Суфіксне дерево», частотність, коефіцієнт взаємної інформації.

#### Summary

In this paper by creating a research corpus we made an attempt to detect collocations in Ukrainian prose of different periods. We explored the features and usage of collocation in artistic texts of male and female authors. For this analysis sub-corpus were created: corpus of texts from 1961 to 1990, this sub-corpus contains (about 710 thousands tokens) male and female authors; corpus texts of 2000s. (about 927 thousands tokens): male and female authors. Using the material of the research corpus we made an attempt to processes methods of collocation (n-grams) analysis. For the analysis we applied algorithm "Suffix tree" of text distribution into clusters, or so-called n-grams. In this research the total array of texts was divided into 2, 3, 4, 5, 6 grams. Quantitative analysis was applied. To detect collocation we also applied coefficient of mutual information. We made conclusions about particularities of collocation usage in "male" and "female" texts of different time periods.

**Key words:** collocation, n-gram, corpus of texts, prose, algorithm of "Suffix tree", frequency, coefficient of mutual information.