

DOI: 10.18524/2307-4558.2021.35.237835

УДК 81'286'374-13:004

**ГРОМКО Тетяна Василівна,**

кандидат філологічних наук, доцент кафедри української мови Центральноукраїнського державного педагогічного університету імені Володимира Винниченка; вул. Шевченка, 1, м. Кропивницький, 25006; тел.: +38 050 0310222; e-mail: hromkot@gmail.com; ORCID ID: 0000-0002-4661-4302

**ТЕЗАУРУС ЯК ЛІНГВОГРАФІЧНИЙ ОПИС ГОВІРКИ: ТЕОРЕТИЧНІ ЗАСАДИ НАУКОВО-ІНФОРМАЦІЙНОЇ БАЗИ**

**Анотація.** *Метою* запропонованої статті є опис та теоретичне осмислення тезаурусу говірки як лінгвографічного продукту, що маючи свою специфіку, становить емпіричне підґрунтя доказової лінгвістики на сучасному її рівні розвитку. *Об'єктом* є тезаурус говірки, який представляє її лінгварій як мовно-мовленнєву систему. *Предмет* дослідження — теоретичні та функційні засади і параметризація тезаурусу говірки як лінгвографічний об'єкт її опису як мовної системи. *Результатом* дослідження є обґрунтування тезаурусу говірки як цілісного інтегрованого продукту у всьому обсязі його лінгвістичного інформаційного наповнення та дотримання вказаних теоретичних настанов; як лексико-граматикографічного комплексу, що має спеціально-наукове лінгвістичне призначення; проміжною ланкою між традиційною картотекою і комп'ютерним корпусом; визначення тезаурусу з усіма його перевагами щодо традиційного лексикографічного представлення в діалектологіїтощо. **Тезаурус** — лінгвографічне видання або електронний ресурс, що містить слова (словосполучення, фразеологізми, паремії тощо), розташовані в алфавітному порядку (з відображенням лексикографічних, граматикографічних, фразеологічних та інших одиниць), що пояснює значення описуваних мовно-мовленнєвих одиниць та надає іншу лінгвістичну інформацію або повідомляє відомості про істот і предмети, які позначаються ними (онімографічні одиниці), а також наводить контексти їх вживання в мовленні певної мовної системи. **Висновки.** Тезаурус говірки має величезний потенціал як нова технологічна можливість опису лінгварію говірки як мовної системи. Як лінгвографічний об'єкт, він спроможний вирішити широкий спектр лінгвістичних завдань в удосконаленні й оптимізації науково-інформаційного забезпечення лінгвістичних досліджень.

**Ключові слова:** лінгвографія, тезаурус, лінгварій, говірка, говірковий дискурс, текстотека, мовна одиниця, мовленнєва одиниця, лексико- і граматикографічний комплекс, параметризація.

**Формулювання проблеми.** Одним із сучасних науково-інформаційних засобів, що спроможний вирішити проблеми лінгвістичної науки, виступає лінгвографія. Лінгвографічні методи і прийоми орієнтують наукову емпірію на об'єктивне відображення мовних даних, на досягнення сучасного рівня технологій дослідження та створення бази для подальшого розвитку лінгвістики в колі не лише гуманітарних, а точних наук. Методологічна цінність лінгвографічних технологій полягає у забезпеченні лінгвальної доказовості через емпіричний матеріал досліджень, що й визначає сенс обробки мовних даних та специфічний інструментарій опису. Так, сьогодишнє лінгвістичне студіювання вже не уявляється без значного масиву різнохарактерних матеріалів та значних інформаційних можливостей й створенні й експлуатації багатоаспектних лінгвістичних баз даних. Лінгвісти-науковці вкрай зацікавлені в оптимізації своєї діяльності, у підвищенні її результативності та ефективності в будь-якій галузі мовознавства, і діалектології зокрема.

Найдосконалішою оптимізацією наукової діяльності мовознавців є проведення досліджень за актуальним нині лінгвографічним засобом — корпусним підходом. За відсутності такого, незважаючи на трудомісткість процесу, рутинну технічну працю, поза увагою науки на повинна залишатися повнота інформаційної забезпеченості дослідження, яку можуть гарантувати інші лінгвографічні техніки: уведення в науковий обіг нових матеріалів, оперативне отримання необхідних зафіксованих (чи експериментальних) даних в систематизованому і наочному вигляді, розширення можливостей обробки, класифікації матеріалу за сучасним лінгвістичним інструментарієм.

**Зв'язок з попередніми дослідженнями.** Теоретичне осмислення лінгвографії, використовуваної лінгвістами як у практиці укладання словників, що є найдавнішим способом і фіксації мовних одиниць, і лексикографування, сьогодні має мало прикладів опису мовленнєвих одиниць. Лінгвографічний опис (опис певної мовної системи) у вигляді зводу мовно-мовленнєвих одиниць уперше пропонується й науково інтерпретується авторкою статті у вигляді тезаурусу говірки. Досі не обґрунтоване місце цієї лінгвографічної праці у класифікаційній системі словників (за типологією, запропонованою Л. В. Щербою [8, с. 305]). Тому теоретичне осмислення тезаурусу говірки базується на авторових розвідках [2; 4] і представлений його як лінгвографічної технології дескрипції говірки [3; 5]. Теорія лінгвографії, категоріально розроблена свого часу на матеріалі російсько-татарського корпусного дослідження [1], може виступати як засаднича для опису мовно-мовленнєвих одиниць в межах тезаурусу.

**Формулювання завдань.** Метою запропонованої статті є опис та теоретичне осмислення тезаурусу говірки як лінгвографічного продукту, що маючи свою специфіку, становить емпіричне підґрунтя доказової лінгвістики на сучасному її рівні розвитку. Об'єктом є тезаурус говірки, який представляє її лінгварій як мовно-мовленнєву систему. Предмет дослідження — ТГ як цілісний інтегрований продукт у всьому обсязі його лінгвістичного інформаційного наповнення та дотримання вказаних

теоретичних настанов; ТГ як лексико- та граматикографічний комплекс, що має спеціально-наукове лінгвістичне призначення; ТГ, з усіма його перевагами щодо традиційного лексикографічного представлення в діалектології; ТГ проміжна ланка між традиційною картотекою і комп'ютерним корпусом тощо. Сучасні тенденції лінгвістики мають вектор закріплення у певній словниковій формі одиниць всіх рівнів мовної структури. Методологія тезауруса певної мовної системи — метамовний інструмент у процесі інтеграції теорії і практики лінгвістичної науки, щодо ТГ є екстраполяцією на говірку шаблону опису окремого лінгварію та апробацією відповідної метамови лінгвістики.

**Виклад основного матеріалу.** Лінгвографія — це 1) міждисциплінарна галузь мовознавства, теорія і практика створення словників, мовних довідників; підрозділами лінгвографії є лексикографія, фразеографія, морфемографія, пареміографія та ін.; 2) сукупність словників (мовних довідників). Лінгвографічне підґрунтя може репрезентувати рівень доказовості лінгвістичного дослідження, оскільки спроможний вирішити проблему наукової інформатизації мовознавства з урахуванням надбань і надання нових даних для його розвитку.

У руслі цього напрямку базою лінгвістичного дослідження є створення металінгвографічного звуку, який інвентаризує емпіричний матеріал, що визначається нами як тезаурус певної мовної системи. Тезаурус є повним словником окремого лінгварію. Тезаурус як мегадовідник дає користувачеві-лінгвісту широкі можливості для пошуку і відбору одиниць, зафіксованих в ньому, з урахуванням різних параметрів. Як основне лінгвістичне джерело, що може і виконати такі запити, і як науково-інформаційний потенціал підтверджує свій статус «обличчя науки» [6] та «лінгвістичного кисню» (*образне утворення наше*. — Т. Г.) є словники. Їх особливість в інформаційному забезпеченні лінгвістики полягає у поліаспектності та поліперентації різного роду мовних довідкових систем. Саме лінгвістична регламентація, параметризація зафіксованих мовних даних [7] дозволяє накопичувати, оптимально структурувати й оперативно видавати для користування лінгвісту-досліднику різноманітну інформацію. Таким сьогодні може послужити своєрідний повний моноговірковий словник — тезаурус, що довідково представляє лінгварій говірки як мовної системи. Лінгвографічні технології, що репрезентують теоретичні засади укладання такого типу словника, актуалізуються на тезаурусі говірки (ТГ).

Загалом ТГ має безперечні переваги як лінгвографічний продукт у порівнянні з традиційними лексикографічними матеріалами. Вони полягають у комплексності реєстрації (інвентаризації) зафіксованих у говірці *мовних і мовленнєвих одиниць*, що забезпечує інформаційні потреби не лише лексикології, а й відображення мовленнєвих даних, що важливе для граматики текстів, колоквіалістики, усно-розмовної дискурсології, контактології тощо. Маючи мовно-мовленнєву специфіку архітектури, як лексикографічна технологія тезаурус говірки є основою інформаційного компонента мовної системи, що як довідник відображає сучасні можливості лінгвографічної системи і в перспективі слугуватиме практиці укладання цілого ряду аспектних словників (а то й стане базою для укладання комп'ютерного корпусу).

За теоретичними засадами сучасної лінгвографії у прикладному моноговірковому дослідженні ТГ забезпечує створення інформаційного компоненту її дескрибування як мовної системи, а також удосконалення та оптимізації лінгвографічної діяльності. Останні як усно-розмовна форма представлення організації інформаційних фондів переглянуті та скориговані щодо діалектологічної науки, яка вивчає не лише мовні, а й мовленнєві факти, явища, різні за характером відображення. Так, лінгвографічна інвентаризація (реєстрація) матеріалів орієнтована саме на дескрипцію, а не на монографічний комплексний опис усіх мовних одиниць говірколінгварію.

Тезаурусна форма подачі матеріалу говіркового дискурсу вважається нами найоптимальнішою для оперативного отримання лексикографічної інформації про говірку [3, с. 115]. Слово як об'єкт лінгвістичного опису та словоцентризм в основі лексикологічних досліджень сьогодні не задовольняють діалектологів, адже усно-розмовна природа діалектного мовлення (а не лише діалектна лексика) потребує особливого типу словників або ж граматично-контекстних довідників, одиницею опису яких має бути граматична інформація, оскільки, на думку Ю. М. Караулова, «будь-які дані про мову можуть бути представлені в лексикографічній формі» [6]. При такому лексикографуванні об'єктом опису можуть ставати фразеологізми, паремії, інші надслівні мовні одиниці і навіть статті-довідки фонетичних експериментів. Звідси «незручності у використанні терміна «лексикографія» і протиріччя внутрішньої форми призвели до утворення нових термінів — *морфемографія, фразеографія, пареміографія* [1]. «Якщо мати на увазі, що існує і традиційне розуміння цього терміна (словник лексики), то *лексикографія* починає виступати як гіперонім і як гіпонім» [Ibid]. Протиріччя у вживанні цього терміна породило термін «лінгвографія» (структурно — *лінгва, мова* (різні мовні одиниці) як об'єкт опису + *графія*), започаткований ученими Казанського університету у 80-х роках ХХ століття, та похідні від нього *лінгвографування, лінгвографічний, лінгвограф*.

**Тезаурус** — лінгвографічне видання або електронний ресурс, яке містить слова (словосполучення, фразеологізми, паремії тощо), розташовані в алфавітному порядку (звідображенням лексикографічних, граматикографічних, фразеографічних та інших одиниць), що пояснює значення описуваних мовно-мовленнєвих одиниць та надає іншу лінгвістичну інформацію або повідомляє відомості про істот та предмети, що позначаються ними (онімографічні одиниці), а також наводить контексти їх вживання в мовленні певної мовної системи.

Як лінгвографічна технологія тезаурус дозволяє скоригувати деякі традиційні лексикографічні положення у наступних додаткових твердженнях. По-перше, це новий тип сучасного об'ємного словника, що передбачає не тільки книжний формат, а й електронну версію, яку можна кваліфікувати як звід електронних лінгвографічних карток, що орієнтований на корпус і зручний для побудови мовних баз даних, і цим самим відображає інформаційні потреби сучасного рівня розвитку науки і техніки. По-друге, упорядкування одиниць в тезаурусі має специфіку параметризації відповідно до їх мовно-категорійного і мовленнєвого сенсу. По-третє, історично-лінгвістичний аспект тезауруса визначають синхронічний стан описуваного мовного об'єкта. По-четверте, відображення лінгварію певного мовного об'єкта, на противагу традиційному повному словнику, в тезаурусі як довіднику впорядкованих мовних і мовленнєвих одиниць відбувається з відповідними до поточного часу їх характеристиками на основі текстів. По-п'яте, тезаурус, орієнтований на поліаспектні дослідження, що потребують багатостороннього аналізу різностороннього, різнорівневого матеріалу, уже містить первісну обробку (інвентаризацію) мовних даних (форма, зміст, особливості функціонування) для потреб науково-інформаційного забезпечення лінгвістики. По-шосте, реєстрація та інвентаризація мовно-мовленнєвих одиниць у тезаурусі є своєрідним конкордансом (з усіма характеристиками та особливостями вживання в джерелі — усному мовленні) певної мовної системи з повним охопленням лінгвального матеріалу. По-сьоме, тезаурус містить більшу кількість і менші обмеження наукової лінгвістичної інформації у порівнянні з лінійно вибудованими традиційними лексикографічними мовними описами.

Таким чином, ТГ може розглядатися як цілісний інтегрований продукт у всьому обсязі його лінгвістичного інформаційного наповнення та дотримання вказаних теоретичних настанов. Виступаючи своєрідним лексикографічним комплексом [5, с. 12], протиставляється словникам як похідним від нього аспектним «підсловникам», доповнюється мовленнєвими одиницями та має спеціально-наукове лінгвістичне призначення. При цьому переваги означуваного нами тезауруса щодо традиційного лексикографічного (чи граматичного, ідіоматичного) представлення очевидні.

На стадії електронних карток тезаурус не дозволяє лінгвістам зручностей у маніпулюванні інформацією як джерелом емпіричного матеріалу подібно до комп'ютерних мовних словників, до того ж, робота з корпусами вимагає вже на етапі селекції матеріалу виробляти відповідну стратегію пошуку. Однак, будучи проміжною ланкою між традиційним і корпусним лексикографічним джерелом лінгвістичних досліджень, тезаурус уже своїм об'ємом вокабуляра є надійним інструментом сучасних наукових пошуків. Відповідно організований та інформатизований, ТГ спроможний забезпечити науковців належними інформаційними можливостями, які навіть не можуть іти в порівняння з традиційними. Такий тезаурус слугує надійним емпіричним підґрунтям доказової лінгвістики на сучасному її рівні розвитку.

Однією з тенденцій сучасного етапу розвитку мовознавства є прагнення представити лінгвістичні матеріали у вигляді довідників, словників, тезаурусів, глосаріїв, яке претендує на закріплення у певній словниковій формі результатів вивчення всіх рівнів мовної структури. У цьому напрямку методологія укладання тезауруса може виступати метамовним інструментом у процесі інтеграції теорії і практики лінгвістичної науки. Запропонований нами досвід створення моноговіркового тезауруса є лише екстраполяцією на говірку як мовну систему і своїми підходами до регламентації, параметризації та лінгвістичної інтерпретації претендує на своєрідний «шаблон опису лінгварію» певної мовної системи (а не предмет накопичення і передачі інформації) та апробацію відповідної метамови лінгвістики. Як мовний довідник, ТГ спроможний функціонувати не тільки як результат наукового дослідження, присвяченого його теоретичному осмисленню [4, с. 115], не тільки як акумулятор у своєму складі отриманої інформації про говірку як лінгвістичний об'єкт, не тільки як джерело моноговіркових даних для подальших досліджень з метою їх уніфікації, а й як інструментар наукового дослідження для обробки мовних даних, що, власне, й відрізняє тезаурус як лінгвографічний об'єкт нашого часу від теорії й практики словникової справи минулих десятиліть.

Тезаурус як лінгвістичний продукт актуально розширяє т. зв. фактографічний фундамент мовознавчих досліджень через залучення в науковий обіг широкого кола різноманітних матеріалів через різного роду електронні картки для оптимізації роботи із системою мовних даних. Такий науковий інструментарій допомагає у вирішенні цілого ряду нагальних, актуальних дослідницьких завдань лінгвістики (у поглибленому вивченні властивостей мови як системи) та її підгалузі діалектології як науки, що потребує широти фіксації фактологічної бази як опори доказовості досліджень. Тезаурус окремого лінгварію — це сучасний крок до науково-інформаційного забезпечення лінгвістики.

Незважаючи на рутину, відсутність творчості та вимогу високого рівня відповідальності в роботі, формування електронної картотеки, ТГ, як показує аналіз, розширяє межі стаціонарного існування зібраного в текстовому форматі фонду. Як проміжна ланка між традиційною картотекою і комп'ютерним корпусом, ТГ є вторинним після текстової експлікації емпіричного матеріалу і таким, що вже має первісну наукову обробку для описування мовного об'єкта та подальших лінгвістичних наукових досліджень. Він спроможний вирішити чимало проблем, які стоять перед сучасною лінгвографією: поліпшення словникової техніки (метамови, структури, параметризації), підвищення точності й експланаторності та надійності лінгвографованих матеріалів, адаптацію матеріалу



в навчальний процес як ефективний лінгводидактичний засіб, а також підготовку матеріалу для представлення (в паперовому чи електронному форматах) широкому загалу культурно-історичного пласту певного мовного об'єднання. Відтак це визначить і форму ТГ: це довідковий багатотомний продукт чи модуль окремої комп'ютерної розробки лексикографічного характеру.

Вивчення властивостей мовних одиниць ТГ може знайти широке застосування при комплексному дослідженні лінгвістики тексту. Адже текст є однією з найскладніших у плані вивчення лінгвістичних сутностей. Одночасно реалізуючи і формуючи систему мови, текст слугує матеріалом для дослідження як явищ мови, так і мовленнєвих явищ. А при укладанні ТГ текст є як базовим для селекції мовно-мовленнєвих одиниць, так і забезпечує урізноманітнення методів і прийомів наукового аналізу текстотеки у процесі подальшої її експлуатації в дослідженнях для досягнення об'єктивності через текстобазувальне їх науково-інформаційне забезпечення.

У ТГ як у лінгвографічному продукті експліцитний опис певної мовної системи передбачає й аналіз мовленнєвих особливостей тексту, які в межах комп'ютерної технології вивчення текстів лінгвіст К. Р. Галіуллін називає *лінгвограмами* (фонетограмами, лексикограмами, морфограмами та синтаксограмами) [1]. На нашу думку, лише попередня обробка й опис тексту з використанням лінгвографічної технології (укладання тезаурусу) спроможні окреслити й подати цінні в лінгвістичному відношенні характеристики тексту та його елементів. Така наукова апробація текстотеки визначається ступенем лінгвістичного опрацювання мовленнєвих одиниць, описом лінгвографічних параметрів і попередньої наукової інтерпретації ТГ.

Тестотека — це фактологічна база лінгвографії, що через параметризацію позиціонує особливості мовних одиниць та їх функціонування в тексті (детальніше див. [2, с. 31]). У межах ТГ мовно-мовленнєві одиниці інтегруються в єдину систему, водночас збагачуючи ілюстративну базу кожної вокабули через різноманітні контексти. У процесі експлуатації лінгвістом текст може виконувати ряд лінгвографічних (лексикографічних, граматикиграфічних, фонетикографічних) функцій для створення словоформопоказчиків, конкордансів, частотних словників тощо.

Як інвентаризувальний (реєструвальний) за характером, ТГ акумулює дані, що містяться в говірці як окремій мовній системі, цим самим створюючи лінгвографічний її фонд, що в перспективі може виступати підфондом національного чи навіть слов'янського лінгвографічного фонду. Інформаційний потенціал тезаурусу зростає за рахунок включення поліаспектних даних лінгварію та високого рівня їх лінгвістичної деталізації, що може послужити надійною базою похідних від нього аспектних словників.

Оскільки лексикографія є найбільш розвиненим розділом лінгвографії (одиницею опису словників є слово), то ТГ із цього боку можна вважати лексикоцентричним (словоцентричним) словником, що відображає лексикографічні класи. Водночас граматичний його аспект складають словоформи і словосполучення, фразеологізми, речення як одиниці граматикицентричних класів мовної системи. Щодо останніх, то інтерпретація мовленнєвих одиниць у складі означуваного нами тезаурусу має свою специфіку словникового опису таких елементів: у репрезентації вокабули, змісті словникової статті, визначенні у філіації не значень, а функції в мовленні тощо. Мовні особливості описуваних у ТГ одиниць базуються на загальних мовних особливостях, що характеризують слово як одиницю мови, а не одиницю словника: відображення вимови традиційною фонетичною транскрипцією, граматичні поміти, фрагменти словозміни, семантика та ін. Це пов'язується із специфікою відображення слова як лексикографічного класу, що «об'єднує ознаки різних аспектів мови: звукового, змістового, граматичного» [9]. Відображення слів граматикиграфічного класу передбачає ознаки як лексикологічного, так і нелексикологічного характеру: фонетичні, морфологічні, синтаксичні. Загальні лексикографічні особливості тезаурусних мовно-мовленнєвих одиниць відображаються у певним чином параметризованій структурі словникових статей ТГ. При цьому варіантність цих одиниць не відображається в межах опису вокабуляру (у словниковій статті). Залежно від ознак, покладених в основу виділення лексико- чи граматикиграфічного класу слова чи мовленнєвої конструкції, кількість одиниць цього лінгвографічного опису статистично відносна, адже може бути або визначена точно, або ж відносно, тому що кордони слова можуть бути розмиті в окремій граматикиграмі, що входить до однієї чи до кількох вокабул ТГ. Тобто вивчення закономірностей сполучуваності слів потребують актуалізації в рамках окремого аспектного дослідження лінгварію говірки.

**Висновки.** ТГ має величезний потенціал як нова технологічна можливість опису лінгварію говірки. Особливості мовно-мовленнєвого представлення та інвентаризації лінгвографічних одиниць різноманітного типу репрезентують ТГ як багатоаспектну словникову базу та надійну фактологічну основу лінгвістичної науки. Відображення не лише слова, а й його явищ та ознак, що надаються можливостями тезаурусу, є одним із шляхів пізнання та аналізу сутнісних характеристик мовних та мовленнєвих одиниць мови як системи. ТГ як лінгвографічний об'єкт спроможний вирішити увесь зазначений у нашій розвідці спектр завдань по відношенню до опису лінгварію говірки як мовної системита удосконаленню й оптимізації науково-інформаційного забезпечення лінгвістичних досліджень.

## Л і т е р а т у р а

1. Галиуллин К. Р. Русская и татарская лингвография : научно-информационное обеспечение : автореф. дис. ... докт. филол. наук. Казань, 2000. URL : <http://dspace.kpfu.ru/xmlui/handle/net/29457/>.
2. Громко Т. В. Сучасна текстографія діалектоносіїв степової говірки. *Наукові записки. Серія : Філологічні науки (мовознавство)*. Кропивницький : Вид-во «КОД», 2019. Вип. 175. С. 29–33.
3. Громко Т. В. Сучасні лексикографічні технології опису говірки як мовної системи. *Науковий вісник Міжнародного гуманітарного університету*. Одеса : Гельветика, 2021. Вип. 46. Т. 1. С. 114–117.
4. Громко Т. В. Тезаурус говірки як показник її онтологічного статусу. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія : Філологія. Соціальні комунікації*. Київ, 2020. Том 31 (70) № 2. Ч. 1. С. 1–6.
5. Громко Т. В. Моноговірковий лексикографічний комплекс як сучасна лінгвістична технологія. *Мова*. Одеса : Астропринт, 2020. № 34. С. 11–17.
6. Караулов Ю. Н. Методология лингвистического исследования и Машинный фонд русского языка. *Машинный фонд русского языка : идеи и суждения*. Москва : Наука, 1986. С. 13–25.
7. Осьмак Н. А. Лексические единицы повседневной разговорной речи : пути лексикографического описания их функционирования : дис. ... канд. филол. наук. Санкт-Петербург, 2014. 235 с.
8. Щерба Л. В. Предисловие [к русско-французскому словарю]. *Щерба Л. В. Языковая система и речевая деятельность*. Ленинград : Наука, 1974. С. 304–312.
9. Языкознание : Большой энциклопедический словарь. 2-е изд. Москва : Большая российская энциклопедия, 1998. 685 с.

## References

1. Galiullin, K. R. (2000), *Russian and Tatar linguography : scientific and information support* : Author's thesis [Russkaya i tatarskaya lingvografiya : nauchno-informatsionnoye obespecheniye : avtoref. dis. ... doct. filol. nauk], Kazan', available at : <http://dspace.kpfu.ru/xmlui/handle/net/29457/>.
2. Hromko, T. V. (2019), "Modern textography of the subdialects of the steppe dialect", *Proceedings, Series : Philological Sciences (Linguistics)* ["Suchasna tekstohrafiia dialektonosiiiv stepovoi hovirky", *Naukovi zapysky. Filolohichni nauky (movoznnavstvo)*], KOD Publishing house, Kropyvnytskyi, vol. 175, pp. 29–33.
3. Hromko, T. V. (2021), "Modern lexicographic technologies for describing subdialect as a language system", *International Humanitarian University Herald. Philology* ["Suchasni leksykohrafichni tekhnolohii opysu hovirky yak movnoi systemy", *Naukovyy visnyk Mizhnarodnoho humanitarnoho universytetu. Filolohiya*], Helvetyka, Odessa, vol. 46 (1), pp. 114–117.
4. Hromko, T. V. (2020), ["Thesaurus of speech as an indicator of its ontological status", *Scientific notes of V. I. Vernadsky Taurida National University*], Series: "Philology. Journalism" ["Tezaurus hovirky yak pokaznyk yiiv ontolohichnoho statusu", *Vcheni zapysky Tavriys'koho natsional'noho universytetu imeni V. I. Vernads'koho. Filolohiya. Sotsial'ni komunikatsiyi*], Kyiv, vol. 31 (70), issue 2, part 1, pp. 1–6.
5. Hromko, T. V. (2020), "Monosubdialectal lexicographic complex as a modern linguistic technology", *Mova / Language* ["Monohovirkoviy leksykohrafichnyi kompleks yak suchasna lnhvistychna tekhnolohiia", *Mova*], Odessa I. I. Mechnikov National University, Astroprint, Odessa, vol. 34, pp. 11–17.
6. Karaulov, Yu. N. (1986), "Methodology of linguistic research and the Machine fund of the Russian language", *Machine Fund of the Russian Language : Ideas and Judgments* ["Metodologiya lingvisticheskogo issledovaniya i Mashinnyy fond russkogo yazyka", *Mashinnyy fond russkogo yazyka : idei i suzhdeniya*], Nauka Publishing house, Moscow, pp. 13–25.
7. Osmak, N. A. (2014), *Lexical units of everyday colloquial speech : ways of lexicographic description of their functioning* : Thesis [Leksycheskiye edynytsy povsednevnoy razgovornoj rechy : puti leksykograficheskogo opisaniya ikh funktsionirovaniya : dis. ... kand. filol. nauk], St. Petersburg, 235 p.
8. Shcherba, L. V. (1974), "Preface [to the Russian-French dictionary]", *Language system and speech activity* ["Predislovie [k russko-frantsuzskomu slovaryu]", *Yazykovaya sistema i rechevaya deyatel'nost'*], Nauka Publishing house, Leningrad, pp. 304–312.
9. *Linguistics : Large encyclopedic dictionary* (1998) [*Yazykoznaniye, Bol'shoy entsiklopedicheskiy slovar'*], Great Russian Encyclopedia, Moscow, 685 p.

## ГРОМКО Татьяна Васильевна,

кандидат филологических наук, доцент кафедры украинского языка Центральноукраинского государственного педагогического университета имени Владимира Винниченко; ул. Шевченко, 1, г. Кропивницкий, 25006, Украина; тел.: +38 050 0310222; e-mail: hromkot@gmail.com; ORCID ID: 0000-0002-4661-4302

## ТЕЗАУРУС КАК ЛИНГВОГРАФИЧЕСКОЕ ОПИСАНИЕ ГОВОРА: ТЕОРЕТИЧЕСКИЕ ОСНОВЫ НАУЧНО-ИНФОРМАЦИОННОЙ БАЗЫ

**Аннотация.** Целью предлагаемой статьи является описание и теоретическое осмысление тезауруса говора как лингвографического продукта, имеющего свою специфику, составляющего эмпирическое основание доказательной лингвистики на современном её уровне развития. **Объектом** изучения является тезаурус говора, который представляет лингварий этого говора как культурно-речевую систему. **Предмет** исследования — теоретические и функциональные основы и параметризация тезауруса говора как лингвографического объекта описания как языковой системы. **Результатом** исследования является обоснование тезауруса говора как целостного интегрированного продукта во всём объёме его лингвистического информационного наполнения при соблюдении указанных теоретических установок;

как лексико- и грамматикографического комплекса, имеющего специально-научное лингвистическое назначение; как промежуточное звено между традиционной картогекой и компьютерным корпусом; определение тезауруса со всеми его преимуществами относительно традиционного лексикографического представления в диалектологии и тому подобное. *Тезаурус* — лингвографическое издание или электронный ресурс, который содержит слова (словосочетания, фразеологизмы, поговорки и т. д.), расположенные в алфавитном порядке (с отображением лексикографических, грамматикографических, фразеологических и других единиц), что объясняет значение описываемых культурно-речевых единиц и предоставляет другую лингвистическую информацию или сообщает сведения о существах и предметах, обозначаемых ими (онимографические единицы), а также приводит контексты их употребления в речи определённой языковой системы. **Выводы:** тезаурус говора имеет огромный потенциал как новая технологическая возможность описания лингвария говора как языковой системы. Как лингвографический объект, он способен решить широкий спектр лингвистических задач в совершенствовании и оптимизации научно-информационного обеспечения лингвистических исследований.

**Ключевые слова:** лингвография, тезаурус, лингварий, говор, диалектный дискурс, текстотека, языковая единица, речевая единица, лексико- и грамматикографический комплекс, параметризация.

**Tatyana V. HRONKO,**

Candidate of Philology (Ph.D.), Docent of the Ukrainian Language Department, Volodymyr Vynnychenko Central Ukrainian State Pedagogical University; 1 Shevchenko st., Kropyvnytsky, 25006, Ukraine; tel.: +38 050 0310222; e-mail: hromkot@gmail.com; ORCID ID: 0000-0002-4661-4302

### THE SAURUS AS A LINGUOGRAPHIC DESCRIPTION OF A SPEAK: THEORETICAL GROUNDS OF SCIENTIFIC AND INFORMATION BASE

**Summary.** The *purpose* of this article is to describe and theoretically comprehend the thesaurus of the dialect as a linguographic product, having its own specifics, constitutes the empirical basis of evidence-based linguistics at its present level of development. The *object* is a dialect thesaurus, which presents its linguistic as a cultural-speech system. The *subject* of the research is the theoretical and functional foundations and parametrization of the dialect thesaurus as a linguistic object of its description as a linguistic system. The *result* of the research is the substantiation of the dialect thesaurus as a holistic integrated product in the entire volume of its linguistic information content and compliance with the indicated theoretical principles; as a lexical and grammatical complex, has a special scientific linguistic purpose; an intermediate link between a traditional filing cabinet and a computer case; definitions of the thesaurus with all its advantages relative to the traditional lexicographic representation in dialectology and the like. *Thesaurus* is a linguistic publication or electronic resource that contains words (phrases, phraseological units, proverbs, etc.) arranged in alphabetical order (displaying lexicographic, grammatical, phraseographic and other units), which explains the meaning of the described cultural and speech units and provides other linguistic information or reports information about the creatures and objects designated by them (onymographic units), and also gives the contexts of their use in the speech of a particular language system. **Conclusions:** The dialect thesaurus has enormous potential as a new technological opportunity to describe the lingual dialect as a linguistic system. As a linguographic object, it is able to solve a wide range of linguistic problems in improving and optimizing the scientific and information support of linguistic research.

**Key words:** linguography, thesaurus, linguistic, subdialect, subdialect discourse, text-library, linguistic unit, speech unit, lexical and grammatical complex, parametrization.

*Стаття отримано 4.05.2021 р.*

DOI: 10.18524/2307-4558.2021.35.237840

УДК 811.163.2'271.12'282.6(560.11)

**ИЛИЕВ Иван Георгиев,**

доктор по български език, доцент в Катедра по хуманитарни науки, Пловдивски университет, Филиал Кърджали; бул. „Беломорски“ 26, Кърджали — 6600, България; тел.: +359886787696; e-mail: ivan\_iliev20002000@yahoo.com; ORCID ID: 0000-0002-2627-972X

### ЗА НАЧАЛОТО НА БЪЛГАРСКАТА КНИЖОВНА НОРМА В РЕПУБЛИКА ТУРЦИЯ

**Анотация.** *Обект* на изследване в статията са първите прояви на български книжовен език в Република Турция, дело на българомохамеданския автор Хасан Уйгун, издател по професия. Това е много важна тема, защото нищо не се знае за съществуването на литературни творби на диалектен български в Турция. *Целта* е читателите да се запознаят с неговите прозаични и поетически творби (разказа *Кървавочервени* и стихотворението *Аз съм помак*), писани на български диалект, и с правописа, използван в тях. Използван е филологически *метод*, базиран на текстови анализ на произведенията. Статията е сериозно *научно постижение* и *резултатът* е разкриването на нови данни. *Практическата стойност* на работата е, че тя представлява първи опит за изнасяне на информация за подобна дейност от представител на българското малцинство в Турция, въпреки че досега са съществували опити за подобни прояви