

ВСТУП

У сучасній лексикографічній теорії та практиці останнім часом особливо актуалізувалися три важливі проблеми.

Першою з них є проблема оновлення фундаментальних лексиконів. Адже швидкий темп сучасного життя супроводжують постійні зміни в мові, отже словникова продукція повинна оновлюватися відповідно до вимог часу. Загальновизнано, що великі словники – це, як правило, взірцеві лексикографічні об'єкти, створені значними колективами висококваліфікованих фахівців. Вони подають яскраві зразки лінгвістичної думки певного історичного періоду, що несуть велику кількість явно і неявно виражених рефлексій тієї чи іншої епохи, в них «укарбовано» історичну пам'ять народу та його мовну свідомість у певний період. Поряд із цим такі лексикографічні твори мають інваріантну цінність, виражену в їхніх лексикографічних структурах, і подають новим генераціям лексикографів нетлінні зразки способів лексикографування мовних явищ, приклади розв'язання лексикографічними засобами різноманітних лінгвістично-когнітивних завдань. Тому при завершенні певної епохи їх не варто здавати в архів, але доцільно повною мірою використати найцінніше, що в них є. Отже, зі сказаного випливає завдання оновлення фундаментальних лексикографічних праць – модифікація їх змісту і структури та приведення у відповідність з новими реаліями. Зрозуміло, що серед усіх лексикографічних праць великі багатотомні тлумачні словники національних мов посідають чільне місце. Однак при їх оновленні виникає ціла низка конкретних питань – починаючи від визначення принципів укладання нового лексикону «поверх» старого і закінчуючи технологією та організацією виконання такого лексикографічного проекту. При цьому технологічні

проблеми все помітніше пересуваються на передній план актуальності, що зумовлено досить стисненими термінами, які відводяться на модернізацію лексику, й необхідністю подання кінцевого результату виконаної роботи у різних, «неізоморфних» між собою формах (традиційній книжковій, у вигляді лазерного диска, сайту в Інтернеті, резидентного словника в комп'ютерних системах опрацювання текстів, електронного семантичного аналізатора тощо).

Друга проблема пов'язана із необхідністю переведення лексикографічної спадщини людства у цифрову форму з метою нагромадження загальнокультурного потенціалу, а також для використання в системах автоматичного опрацювання інформації. Для цього абсолютно недостатньо просто сканувати тексти відповідних словників, хоча й це не така проста задача, як може здатися з першого погляду. Відомо, що метамова багатьох словників містить цілий ряд нестандартних символів, ідентифікувати які складно. Крім того, існуючі програми оптичного розпізнавання текстів далеко не бездоганні й роблять чимало помилок у процесі їх застосування. Нарешті, робота з великими (передусім багатотомними тлумачними) словниками, переведеними у комп'ютерні текстові формати, є дуже неефективною і потребує їх конверсії у формати лексикографічних баз даних, що є спеціальним завданням, не відомим класичній лексикографії.

Третя проблема пов'язана із застосуванням словників у формуванні лінгвістичних компонент концептографічних систем експлікації знань (типу онтологій) та використанні їх у засобах knowledge mining¹. Це потребує не лише великих словникових масивів, але й спонукає до віднаходження у традиційних словникових текстах прихованих семантичних структур.

Очевидно, що ефективне вирішення відзначених проблем потребує певної узагальнювальної теоретичної схеми, спроможної надати уніфіковані засоби до їх розв'язання. Таким теоретичним інструментом автори вважають створену в Українському мовно-інформаційному фонді в 90-х роках минулого століття теорію лексикографічних систем [276], одним із найяскравіших практичних застосувань якої стало її використання для побудови формальної структури 11-томного академічного тлумачного Словника української мови (СУМ-11) [85]. Ця робота дозволила не лише глибше усвідомити лінгвістичні ефекти, «зашифровані» в структурі цього Словника, але й скласти підґрунтя для побудови в автоматичному режимі розвине-

¹ Видобування знань (англ.).

ної і складної комп'ютерної бази даних Словника української мови [289]. Наскільки нам відомо, до цього у світовій комп'ютерній лексикографії не вдавалося виконати аналогічну роботу для лексиконів такого масштабу і складності.

Створення лексикографічної бази даних на основі 11-томного Словника української мови стало реальною передумовою до укладання нової – 20-томної – версії академічного тлумачного Словника української мови (СУМ-20). Оновлення фундаментального лексикону 11-томного Словника української мови на рубежі тисячоліть виявилось не лише актуальним мовознавчим завданням, але й узагалі нагальною потребою українського суспільства. Власне, 20-томний тлумачний Словник є одним із центральних проектів програми створення Національної словникової бази, ініційованої Указом Президента України від 7 серпня 1999 р. № 967. Як окреме завдання «Створення фундаментальної академічної багатотомної лексикографічної системи “Словник української мови”» проект є в розпорядженні Кабінету Міністрів України від 22 листопада 2000 р. № 464-р «Про першочергові завдання із створення національної словникової бази». Відповідно до цього розпорядження призначено й координатора роботи – Український мовно-інформаційний фонд НАН України. Постановою Президії НАН України від 16.02.2001 р. № 45 визначено головного наукового редактора проекту «Словник української мови» – академіка В. М. Русанівського, під керівництвом і у співробітництві з яким одним із авторів цієї книги (В. А. Широковим) було розроблено лінгвістичну концепцію Словника [241], що протягом 2001–2006 років була втілена у його першій редакції. Друга редакція СУМ-20 здійснювалася протягом 2007–2010 років. Її першим результатом є лексикографічна база даних «СУМ-20» та перший і другий паперові томи цього Словника, які незабаром вийдуть друком у видавництві «Наукова думка».

Основне завдання пропонованої читачеві книги – донести зміст лексикографічної концепції, технології та організації праці, які склали підґрунтя для створення СУМ-20, і на цій базі й на цьому прикладі подати нарис лінгвістичних принципів та технологічних засад створення великих тлумачних лексиконів. Важливим питанням, яке до початку нашої праці навіть не ставилося у практичному ключі, є наступне: наскільки новий словник відрізняється від свого прототипу і чи насправді він є суттєво новим лексикографічним продуктом? Відповідь на нього порушує важливі для словникарства проблеми лінгвостатистики та лексикографічної компара-

тивістики, які зі зрозумілих причин не можуть бути розв'язані без застосування комп'ютерних методів.

Усі вищевикладені аспекти тією чи іншою мірою представлені у цій книзі.

На жаль, В. М. Русанівський не дожив до виходу бодай першого з томів нового Словника української мови й навіть не встиг написати вступної статті. Віталій Макарович написав лише початок Вступу до Словника, з якого ми й починаємо виклад.

«Нація усвідомлює себе тоді, коли вона створює порівняно повний словник своєї мови. Українська нація у 1970–1980 рр. дістала таку працю – “Словник української мови” в 11 томах (далі – СУМ-11; для 20-томника вживаємо аббревіатуру СУМ-20). У ньому зафіксовано, описано й розтлумачено 134 тисячі слів нашої мови. Це приблизно стільки ж, скільки на той час описувалося і в лексиконах не тільки слов'янських, а й інших європейських мов. Проте мова не стоїть на місці, як і сам народ. Україна наприкінці ХХ ст. нарешті здобула незалежність і разом із нею – суверенітет власної мови. Виявилось, що, по-перше, її словниковий склад значно поповнився новими словами; по-друге, відійшла на задній план лексика з так званої радянської “новомови”; по-третє, змінилася семантика багатьох слів. Тому постала потреба у новому українському лексиконі. Саме таким і став “Словник української мови” в 20 томах, створений у рамках програми Національної словникової бази України.

Для систематичного опису лексики потрібне її якомога повніше зібрання. Досі цьому служили лексичні картотеки, які створювалися упродовж багатьох десятиліть. Такою є шестимільйонна картотека, зібрана в Інституті мовознавства ім. О. О. Потебні, а нині підпорядкована Інституті української мови Національної академії наук України. Саме на цій базі укладався виданий у другій половині ХХ ст. “Словник української мови”. На цьому її можливості не були вичерпані: вона ще може багато в чому прислужитися дослідникам української мови. Але для нового словника тлумачного типу цієї картотеки вже замало, тим більше, що укладений на її основі “Словник української мови” 1970–1980 рр. повністю входить до складу нового, 20-томного тлумачного словника. Доповнюють його матеріали нової, електронної картотеки – лінгвістичного корпусу, створеного в Українському мовно-інформаційному фонді Національної академії наук України, які на момент укладання Словника нараховують понад 43

мільйони слововживань². Джерелом інформації для Словника став також Інтернет.

Завдання нового Словника, який укладено в паперовому й електронно-му варіантах, – з максимальною увагою представити лексичне, а головне, – семантичне багатство сучасної української мови, яке творилося від її зародження та відображене в письмових джерелах XIX, XX і XXI ст. Кількість цих джерел, порівняно з лексичною картотекою, зросла більше ніж удвоє.

Для цього в лінгвістичному корпусі зібрано нові, а також не нові, але наново відкриті джерела: твори заборонених раніше письменників, сучасних творців художнього слова, зарубіжних українськомовних письменників, сучасна гуманітарна і науково-технічна література. Беручи до уваги, що в наявних текстах, особливо в публіцистиці, все ще трапляються слова з колишньої радянської “нової мови” (на зразок *ударник*, *стахановець* та ін.), вони залишаються в словнику з ремаркою *заст.* Вилучаються тільки ті, які в новому суспільстві втратили всякий сенс: *автоброньовик*, *антиленінський* (і похідні).

Те саме й зі словосполученнями.

Докорінно переглянуто ілюстрації. Знято ті, що походили з одіозних джерел. Замінено цитати з творів того самого письменника, якщо попередні мали заідеологізований характер, залучено десятки тисяч нових ілюстрацій із творів, наявних у лінгвістичному корпусі та інших джерелах. Значно скорочено кількість перекладних творів (залишено лише світову класику) і кількість перекладачів. Цю місію у виданні виконують тільки М. Рильський, Борис Тен і М. Лукаш³. Звичайно, зовсім знято як джерело ілюстрацій усе, що стосувалося документів КПРС і КПУ.

В 11-томному Словнику занадто щедро практикувалося уживання загальномовних і стилістичних ремарок на зразок *книжн.*, *вульг.*, *ірон.*, *лайл.*, *фам.*, *поет.*, *нар.-поет.*, *спец.*, *розм.*, *жарг.*, *заст.*, *уроч.*, *діал.* Вони

² Електронну картотеку, згідно із загальноприйнятою зараз термінологією, ми називатимемо лінгвістичним корпусом. Теорія та технологія лінгвістичних корпусів викладена в колективній монографії «Корпусна лінгвістика» В. А. Шарокова, О. В. Бугакова, Т. О. Грязнухіної та ін. [283]. Обсяг Українського національного лінгвістичного корпусу, який функціонує в Українському мовно-інформаційному фонді НАН України, зараз сягає понад 64 млн слововживань.

³ У СУМ-20, крім згаданих В. М. Русанівським перекладних творів М. Рильського, Бориса Тена і М. Лукаша, подано ілюстрації з перекладів М. Бажана, М. Вороного, М. Зерова, Г. Кочура, С. Масляка, М. Терещенка і В. Мисика.

в новому варіанті зберігаються, але їхня кількість зменшується, особливо ремарки *діал.*, яка обмежує уживання відомих слів у звичайних текстах. Що ж до ремарки *заст.*, то частота її вживання зростає у зв'язку з переходом у пасив активних слів попереднього партійно-державного, тоталітарного устрою».

Цими рядками завершується частина Вступу, яку встиг написати В. М. Русанівський. Подальший виклад теоретичних положень і практичних засад укладання СУМ-20 (теорія лексикографічної структури, будова словникової статті та Словника в цілому, його відмінності від 11-томника і т. ін.) був виконаний уже після відходу Віталія Макаровича з цього життя.

Вважаємо за необхідне згадати ще одного з натхненників нового Словника української мови – Марію Михайлівну Пешчак, яка пішла з життя вже під час написання цієї книги. Саме з її легкої руки та за її впливу започатковано фундаментальне дослідження структури «Словника української мови», що привело до побудови теорії лексикографічних систем та значного прогресу в галузі комп'ютерної лексикографії в Україні. Вона й стала одним із організаторів даного проєкту на його початковій стадії: переконувала у необхідності нової версії Словника, проводила співбесіди з мовознавцями, домовлялася з окремими лексикографами про їхню участь у проєкті й сама виконувала обов'язки наукового редактора та укладача одного з томів. На жаль, і їй не судилося побачити надрукованим перший том нового Словника...

За задумом авторів, Словник складається з двох великих частин, які, відповідно, подають загальну та ономастичну лексику (остання містить географічні назви), що відрізняє 20-томник від 11-томника, у якому географічні назви не подавалися. Зазначена новація вмотивована тим, що ці назви є дуже частотними у мовній практиці, а їх написання і тлумачення не завжди просте. Крім того, вимога лексикографічної повноти та замкненості, яка є обов'язковою для великих лексиконів, спонукає до внесення у реєстр усіх слів, наявних у словникових дефініціях, де трапляються й географічні назви, а вимога системності лексикону вмотивовує подання повномасштабного «географічного» підрозділу Словника. У 20-томнику перші 18 томів містять загальну лексику активного і пасивного лексичних фондів, а географічні назви буде подано в останніх двох томах разом із теоретичними засадами укладання ономастичного масиву Словника, а та-

кож будовою його лексикографічної системи. Друга принципова відмінність нового Словника від його попередника полягає в тому, що він паралельно створюється у двох варіантах – паперовому та електронному (це відзначив і В. М. Русанівський у написаному ним фрагменті вступної статті до Словника). При цьому не варто думати, що електронна форма Словника є лише переведеним у цифрову форму текстом того, що вважається його паперовим оригіналом (якщо такий взагалі існує). Як свідчить досвід Українського мовно-інформаційного фонду, де створено уже десять версій електронного українського словника «Інтегрована лексикографічна система “Словники України”», електронний словник є принципово відмінним від паперового мовно-інформаційним об’єктом. Маючи практично необмежений у порівнянні з паперовими виданнями обсяг, він дає можливість зібрати і представити в єдиному, «зв’язному» лексикографічному продукті значну кількість якісно різнорідних лінгвістичних даних, а також дозволяє експлікувати такі мовні факти, функції та процеси, які просто фізично неможливо подати в паперовому словнику. Нарешті, електронна форма вносить значну новизну в прагматику використання словника, дозволяючи застосовувати його не лише у довідниковому режимі, але і як резидентну систему в контурах опрацювання природномовної інформації.

Слід зазначити, що у процесі вивчення СУМ-11 (ще до початку роботи над новим Словником), а згодом – у процесі укладання СУМ-20 у його авторів виробився погляд на тлумачний словник як на певну «*Summa Lexicographiae*», тобто на такий лексикографічний об’єкт, який «прагне» відтворити у максимальній повноті системних зв’язків якомога більше аспектів лексичної та граматичної семантики, що виявляє словесна матерія. До такого «максималізму» спонукала й форма, у якій втілювалася робота над СУМ-20, особливо – в його другій редакції. Це форма комп’ютерного інструментального комплексу у вигляді віртуальної лексикографічної лабораторії, що стала принципово новою реалією в історії української лексикографії.

Зрозуміло, що задля комплексного, «інтегрованого» висвітлення і лінгвістичних, і структурних, і технологічних аспектів нового Словника (та й лексикографічного опису мови взагалі) необхідна певна загальна концептуальна схема, якою у цій книзі виступає теорія лексикографічних систем, про яку вже згадувалося вище. Тому в монографії подано нарис основних засад цієї теорії. Але перш ніж перейти до її викладу, доцільно розглянути традиційні погляди на тлумачну лексикографію – адже все нове, як відомо,

мо, виростає зі старого. Слід також зауважити, що у вітчизняній лінгвістичній монографістиці практично відсутні ґрунтовні праці, спеціально присвячені комплексному висвітленню лінгвістичних і технологічних проблем сучасної тлумачної лексикографії. Це служить додатковим мотивом для опублікування книги із проблематики тлумачних лексиконів. Разом з тим враховано той факт, що сучасний історичний період розвитку лексикографії певною мірою є перехідним від традиційної, паперової форми до цифрової, електронно-комп'ютерної, орієнтованої на застосування в лінгвістичних технологіях опрацювання інформації та мережових комунікаціях. Такий «проміжний» стан лексикографічної галузі (коли ще не винайдено інтелектуальних автоматів для продукування словникових систем і фігура лексикографа залишається центральною у цих процесах, хоча чимало рутинних, а також деякі інтелектуальні функції лексикографування вже піддаються автоматизації) спонукає до фіксації здобутків цієї галузі, її традицій, рефлексій, технік лексикографування, визначаючи водночас і прагнення висвітлити ті нові тенденції, які диктуються лексикографічними запитами лінгвістичних технологій.

Стосовно призначення цієї книги зауважимо, що вона адресована не лише практичним словникам, яким корисно мати виклад досвіду роботи фахового колективу, що впродовж значного періоду працює над фундаментальним лексикографічним проєктом. Важливим адресатом монографії, на нашу думку, є також широке коло мовознавців та інформаційників — як теоретиків, так і представників прикладної лінгвістики, лінгвістичної технології, когнітивістики та ін., хто цікавиться фундаментальними підвалинами сучасної лексикографії та її зв'язками з різними галузями лінгвістики, логіки та інформатики. Тому значну частину викладу присвячено неklasичним підходам до лексикографічної проблематики, серед яких особливо відзначається концепція семантичних станів і лексикографічних числень, а також питання комп'ютерних технологій лексикографування.

Основний зміст книги складається з шести розділів.

Перший розділ «Тлумачна лексикографія на загальнолінгвістичному тлі» присвячений основним принципам, методам і засадам лексикографічної діяльності та місце і ролі в ній великих тлумачних лексиконів. У цьому ж розділі викладено концепцію семантичних станів та її застосування при визначенні семантичного стану лексеми з метою формалізації процесу побудови словникових дефініцій. Розділ в основному написано В. А. Широковим; у його написанні також брали участь Н. М. Зайка та Л. Л. Шевченко.

У другому розділі викладено теорію лексикографічних систем, яка надає поняттєвий апарат і концептуальні засоби для побудови формальних структур словників будь-яких типів, у тому числі й тлумачних. Розділ написано В. А. Широковим.

Зазначену теорію використано у третьому розділі «Системно-лінгвістичні аспекти Словника української мови» для побудови формальної структури лексикографічної системи цього Словника, виходячи з «перших принципів». У цьому розділі досліджено приховану структуру та симетрію, яку вдалося виявити завдяки застосуванню положень теорії лексикографічних систем при дослідженні словникових статей дієслова та іменника. Зазначену симетрію через проведення обчислювального експерименту вдалося простежити на великому мовному матеріалі, що дає підстави для її екстраполяції на всю мовну систему української мови як певний її системно-структурний закон. Третій розділ написано В. А. Широковим, Н. М. Заїкою та О. О. Погрібною.

У четвертому розділі «Головні засади укладання СУМ-20» викладено загальнолінгвістичні принципи створення нового Словника, критерії добору нової лексики, правописні засади тощо. Розділ написано В. А. Широковим та Л. Л. Шевченко.

П'ятий розділ «Подання мовного матеріалу в СУМ-20» присвячено докладному лінгвістичному описові будови словникових статей СУМА, принципів розроблення словникових дефініцій, лексикографічній характеристиці реєстрових одиниць за частинами мови, розробленню словосполучень різних типів, поданню лексики та фразеології Святого Письма, а також нової та термінологічної лексики. Крім цього, тут викладено систему ремаркування реєстрових одиниць, висвітлено принципи подання текстово-ілюстративного матеріалу, а також описано засади поліграфічного оформлення структури та змісту словникових статей та будову словопокажчика. У написанні розділу взяли участь В. А. Широков, Л. Л. Шевченко, В. М. Білоноженко, Н. М. Заїка, О. О. Погрібна, В. В. Чумак, Г. М. Ярун, Л. В. Туник, О. В. Бугаков, І. В. Шевченко, І. М. Кондрачук, Н. В. Щербакова, М. Ю. Кригін, І. В. Остапова, С. І. Волошина, М. П. Слабковська.

У шостому розділі «Технологічні аспекти проекту "СУМ-20"» викладено системотехнічні засади побудови комп'ютерного інструментального комплексу «Фундаментальна академічна лексикографічна система "Словник української мови"», який було створено в Українському мовно-інформаційному фонді спеціально для реалізації проекту СУМ-20. Тут також

описано будову та функціональні властивості Українського національного лінгвістичного корпусу – основного джерела текстово-ілюстративного фонду СУМ-20. Подано інформацію з лексикографічної статистики щодо цілої низки показників, які наочно демонструють відмінність між СУМ-11 та СУМ-20. Цей розділ написано В. А. Широковим, К. М. Якименком, Н. М. Сидорчук, В. В. Манако, Ю. В. Леонтьєвим.

Розділ 1

ТЛУМАЧНА ЛЕКСИКОГРАФІЯ НА ЗАГАЛЬНОЛІНГВІСТИЧНОМУ ТЛІ

1.1. Означення, функції та загальна структура словника

Філологічна наука за багатовікову історію свого існування і розвитку виробила системні підходи до означення поняття словника. У вузькому розумінні словник (лексикографічний твір) – це сукупність мовних одиниць, розташованих у певному порядку, де розкрито їхнє значення, подано про них різні відомості чи переклад іншою мовою або вміщено інформацію про предмети, явища та факти, які вони позначають.

Більш широке означення словника, яке належить бельгійському лексикографу К. К. Бергу [319], формулюється так: «словник – це систематизований перелік соціалізованих (тобто тих, що стали стабільними в системі мови) мовних форм, узятих зі звичайного мовлення певної мовної спільноти та описаних укладачем так, що підготований читач розуміє зміст кожної окремо взятої мовної форми й отримує інформацію про істотні факти, пов'язані з функціонуванням цієї форми в даній мовній спільноті».

Якщо ж словник розуміти як підсумок дослідницької роботи над якоюсь проблемою, як розв'язок певної лінгвістичної проблеми чи як інструмент розв'язання нових завдань, то поняття «словник» розширюється у напрямі позначення систематизованої сукупності знань про певну проблему [164] і набуває цілком визначених конотацій до когнітологічних структур типу моделей знань.

Теоретична лексикографія подає означення словника як абстрактного мовно-інформаційного об'єкта, визначальними рисами якого слугує певна низка системно-структурних ознак.

Передусім – це членоване розміщення матеріалу, оскільки основною композиційною та комунікативною одиницею словника [149] є відносно самостійний відрізок тексту, що називається словниковою статтею; множина словникових статей і складає основу словника.

У свою чергу, кожна словникова стаття є, як мінімум, двокомпонентною – в ній виділяється реєстрова (ліва) та інтерпретаційна (права) частини. Ліва частина – це, зазвичай, будь-яка одиниця мови, що є об'єктом лексикографування і вноситься до словника. Вона називається ще лексикографічною або реєстровою одиницею. Множина реєстрових одиниць словника складає його реєстр. Структура і зміст правої частини словникової статті залежить від типу словника; тут подається лінгвістичний опис відповідної реєстрової одиниці (її значення, переклад іншими мовами, наголос, лексичні, граматичні характеристики тощо) або міститься різноманітна інформація про предмети, які вона позначає. На склад реєстру та структуру словникової статті, окрім того, дуже впливає кінцеве призначення словника. Зазначимо, що терміни «ліва» та «права» частини словникової статті є певною мірою умовними, оскільки у деяких типах словників спостерігається «чергування» або «перемішування» елементів структур реєстрової та інтерпретаційної частин.

Реєстрові одиниці у словнику пов'язані між собою численними структурно-семантичними зв'язками і в своїй сукупності творять певну систему, яка реалізує задуми та цілі його укладачів. Окрім того, словникові статті будь-якого словника повинні мати – в його межах – тотожні схеми опису однотипних елементів лексикографічних структур. Сукупність правил, прийомів та засобів опису реєстрових одиниць творять метамову словника.

Взагалі метамова є одним із основних понять сучасної логіки та теоретичної лінгвістики, яке використовується при дослідженні мов різних логіко-лінгвістичних числень та природних мов. Метамова – це мова опису, мова «другого порядку», за допомогою якої описуються мови-об'єкти. Для того, щоб відрізнити метамову від мови-об'єкта, її вибудовують як певну «термінологічну систему». З одного боку, вважається, що чим строгіша і чіткіша ця система, тим більше вивченням і, відповідно, науковим є сам об'єкт опису. З іншого, саме у нечіткості та розпливчастості метамови бачать адекватність об'єкта вивчення [233]. За визначенням, наведеним в енциклопедії «Українська мова», метамовою є знакова система, яка використовується для опису природної мови [88]. У лексикографії метамовою словника називають сукупність формально визначених елементів (маркерів), якими марковано елементи словникової структури й лінгвістичні

факти, суттєві з огляду на їх подання у прийнятій укладачами словника моделі предметної галузі, що виступає об'єктом лексикографування.

Важливою характеристикою словника є нормативність – вона фактично слугує (точніше, повинна слугувати) імперативною настановою для лексикографів. Зміст цієї настанови полягає в тому, що при описі кожної лексикографічної одиниці її слід певним чином співвідносити із конкретною формою існування (функціонування) мови, регламентуючи правила її мовленнєвого вживання. Доречно підкреслити, що укладач словника зазвичай лише описує та досліджує мовні процеси, визначає мовні закономірності, проте аж ніяк не може зобов'язувати, вказувати, йти всупереч природному розвитку мови.

Текст словника є строго упорядкованим – реєстрові одиниці у кожному словнику обов'язково подають упорядкованими за якимось критерієм (наприклад, за абеткою).

Нарешті – наявність допоміжного матеріалу. Будь-який словник, окрім словникових статей, містить, хоча б у мінімальній кількості, додаткову інформацію, що допомагає ним користуватися: передмову, критерії впорядкування та опис структури словникових статей, список скорочень, графічні таблиці, джерела укладання тощо.

У структурному аспекті класична теоретична лексикографія розрізняє три взаємопов'язаних рівні ієрархії: макро-, медіо- та мікроструктуру словника [106; 144].

До макроструктури традиційно відноситься все те, що визначає словник як самостійну систему з її внутрішніми зв'язками та багатопланою організацією. І тут теоретичне словникарство вирішує цілу низку проблем. Важливе місце серед них займає створення науково обгрунтованої типології словників; крім того, макроструктурними завданнями є обгрунтування необхідності розроблення конкретного словника, вивчення та відбір джерел до його укладання, визначення загальної структури словника, опрацювання принципів його побудови й укладання, способи творення реєстру та визначення характеру лексики, яку належить включити в словник, визначення можливих перетворень лексичного складу мови (компресії, мінімації) для укладання реєстрів словників різних типів, вироблення метамови словника, уніфікація композиції і апарату посилань однотипних словників для полегшення користувачам переходу від одного словника до іншого, принципи відбору ілюстративного матеріалу, визначення критеріїв упорядкування реєстрових одиниць, відображення у словнику статистичних даних тощо.

До основних проблем медіоструктурного рівня словника відносяться побудова та експлікація різноманітних лінгвістичних відношень між словниковими одиницями. Вони втілюються в об'єднанні реєстрових одиниць у поля та групи на основі морфологічних, семантичних, тематичних, асоціативних та інших ознак.

На мікроструктурному рівні вирішуються проблеми словникової статті: її структура, форми та способи розкриття семантики реєстрових одиниць, ієрархія їхніх значень тощо. Теоретичною лексикографією встановлено такі вимоги до лексикографічного опису як стандартність (єдиний протягом цілого словника опис однотипних явищ), економність (надання переваги – за однакових інших умов – коротшому описові перед довшим), простота (застосування при лексикографічному описі максимально простих та зрозумілих синтаксичних конструкцій і слів, в лексичному значенні яких нема непотрібної двозначності й зайвої розпливчастості), повнота (прагнення вичерпного опису всіх важливих значень і способів вживання реєстрових одиниць).

Існують й інші способи визначення структури словника та рівнів його структури.

Корисним поняттям для визначення структури словникових систем, а також для побудови їхніх класифікаційних схем слугує введене Ю. Н. Карауловим [164] поняття лексикографічного параметра. За його визначенням, лексикографічний параметр – це деякий «квант» лінгвістичної інформації, що може мати самостійний інтерес для користувача, але, як правило, виступає в комбінації з іншими параметрами – «квантами» – і знаходить своє специфічне вираження у словниках; іншими словами – це особливе словникове відображення окремих структурних рис мови.

Для усунення неоднозначності у трактуванні лексикографічних параметрів надалі ті з них, що відносяться до словника в цілому (призначення словника, коло користувачів, спосіб використання, об'єм, кількість мов, входів тощо), називатимемо словниковими параметрами. Лексикографічними параметрами у власному значенні вважатимемо ті, що співвідносяться з лексикографічною одиницею. Наприклад, орфографічний (правописний) параметр, графічна довжина слова, наголос, вимова (орфоепічний параметр), поділ на склади, частиномовна належність тощо.

Лексикографічним параметрам притаманні певні властивості [140; 164], які мають універсальний характер і не залежать від типу словника, а їхні значення характеризуються певною специфікою конкретної мови (наприклад, рід іменників), у деяких мовах певні параметри можуть бути й відсутніми. Кількість лексикографічних параметрів у словниках ва-

ріюється від одного до декількох десятків, причому ці коливання визначаються не лише цільовою установкою словника, але і його орієнтацією на того чи іншого користувача.

Лексикографічний параметр завжди відноситься до лексикографічної одиниці в цілому: це не склад, а поділ на склади, не афікс чи суфікс, а морфемне членування, не окреме словотвірне значення, а комплекс словотвірних відношень або гніздо, не фонема і не звук, а вимова і т. д.

Як віддзеркалення структури всієї мови лексикографічні параметри різняться за своїм обсягом, подекуди вони здаються несумірними один з одним (наприклад, наголос та словотвірне гніздо). Деякі параметри можна виділити досить легко (наприклад, довжина слова), визначення ж змісту інших вимагає попереднього, інколи значного, дослідження (наприклад, семантичний еквівалент слова або його етимологія).

Один і той самий параметр можна задавати різними способами. Окрім того, практично будь-який параметр залежно від кінцевого завдання словника розкривають з різним ступенем глибини та деталізації. Прагнення до деталізації параметрів проявляється в укладанні однопараметричних словників. Можна сказати, що кожен параметр прагне стати окремим словником. Така відцентрова тенденція, як відзначалося, прямо протилежна доцентровій тенденції до універсальності, тобто об'єднання всіх лексикографічних параметрів в одному словнику (те, що ми називаємо «сумою лексикографії»). Ці різноспрямовані тенденції разом творять характерне прагнення лексикографії закріплювати результати мовних досліджень на всіх ділянках лінгвістики у формі словників. Слід зазначити, що збільшення глибини відображення одного параметра обов'язково веде до розширення опису його зв'язків з іншими параметрами мовної структури, до втягнення їх у сферу представлення цього параметра, тобто при досягненні певної глибини в розробці окремого параметра відцентрова тенденція перетворюється на свою протилежність. Багато параметрів виступають у синкретичній формі: акцентологічна інформація подається або в поєднанні з орфографічною, або прив'язується до орфоепічного параметра; поділ на склади передбачає наявність даних про вимову слова і т. ін.

За своєю природою лексикографічні параметри поділяють на дві групи. До однієї з них входять власне мовні, тобто структурогенні параметри — наголос, орфографічний параметр тощо. Їхні значення є дискретними і можливість варіювання розповсюджується лише на способи задання, фіксації у словнику. Другу групу становлять параметри, зміст яких за необхідністю включає і екстралінгвістичний фактор — денотативний, історико-культурний, прагматичний, тобто моменти, вторинні у відношенні

до мови, пов'язані не лише з самою мовою, але з її вивченням, тобто з мовознавством. Ці параметри вбачаються недискретними, характеризуються змінною глибиною наукового розкриття і відображають не власне структурні відношення, а відтворюють процеси – діахронічні, синтагматичні, прагматичні, взаємодії мов, інтерференції тощо. Називатимемо їх лінгвістичними параметрами. Розгляд сучасних словників різних мов не дає підстав для обмеження будь-якої сполучуваності параметрів, конструюванні своєрідного параметрично-лексикографічного простору. Цей висновок підтверджується й історичним розвитком лексикографії в напрямку нарощення числа параметрів у словнику. Характеризуючи параметри цих двох груп, відзначимо, що можна побудувати словники, які містять лише структурогенні параметри, але неможливо створити хоча б скільки-небудь репрезентативні словники з одних лінгвістичних параметрів – останні обов'язково повинні доповнюватись і деякими структурогенними параметрами. З іншого боку, хоча самі параметри можуть сполучуватись у різних комбінаціях, далеко не всі способи їх задання та ступінь глибини виявляються сумісними один з одним. Таким чином, одна зі складностей на шляху побудови універсального словника полягає у встановленні оптимального співвідношення між способами задання структурогенних та ступенем глибини лінгвістичних параметрів у одному словнику. «Геометрична» інтуїція в цьому випадку твердить про знаходження змістових «гіперповерхонь» у параметричному просторі, які експлікують «правильне» співвідношення між різними лексикографічними параметрами.

Практично в усіх словниках вхід здійснюється за одним (рідше – за декількома) параметрами, але обов'язковою умовою є те, що ці параметри лише структурогенні й не можуть бути лінгвістичними. Більше того, число вхідних параметрів виявляється досить обмеженим, тобто далеко не всі структурогенні параметри можуть реально використовуватись як вхідні. В універсальному словнику, де теоретично повинні поєднуватись всі лексикографічні параметри, бажано мати не один, а декілька входів і не лише від основних структурогенних параметрів, але й від ряду лінгвістичних. Зрозуміло, що цю проблему дуже важко, навіть практично неможливо, розв'язати методами традиційної лексикографії.

Лексикографічні параметри закріплюють формальне вираження своїх значень, як правило, в системі словникових ремарок, хоча є параметри (наприклад, дефініція, орфографія), що спеціальних ремарок не мають. Ремарки однакового порядку можуть бути і значеннями одного параметра, допускаючи його більшу чи меншу деталізацію, наприклад, ремарки сти-

лістичні, але можуть бути й різними параметрами, як у випадку системних відношень (синоніми, антоніми тощо).

Традиційний словник обмежений щодо обсягу – реальної фізичної можливості подання лексикографічного матеріалу. А сучасне поліграфічне виробництво є на порядки дорожчим у підготовці багатотомного видання порівняно, скажімо, з випуском лазерного диску, на якому може зберігатися понад 300 тисяч сторінок звичайного тексту, не кажучи вже про інші види інформації (відеофрагменти, звуковий супровід та ін.). Тому очевидно, що прийнятий поділ традиційних словників за обсягом – великий (повний, інтегральний), малий (вибірковий, короткий), середній і тезаурус-скарбниця [232] – не є релевантним для комп'ютерних словників та лексикографічних баз даних, хоча й за традицією вживається у їхніх назвах.

Згадуючи вказані об'єкти у зв'язку з тим, що комп'ютерне лексикографічне моделювання мовних систем з суто філологічної проблеми перетворилося на нагальну потребу створення лінгвістичних компонентів програмного забезпечення, наголосимо на безперечній актуальності відтворення у лінгвістичному забезпеченні якомога більшого числа ефектів природної мови, причому максимально гнучким та універсальним способом. Цим пояснюється справжній бум у створенні комп'ютерних словників, оскільки словник є тим компактним і добре впорядкованим середовищем, яке відносно легко піддається алгоритмізації і може бути ефективно використане у лінгвістичних комп'ютерних компонентах. У зв'язку з такою активністю в галузі лінгвістичних технологій актуалізувалася потреба визначення кола словникових систем, необхідних для комп'ютерних застосувань. Зазначені чинники обумовили також і необхідність створення формальних теорій словників, спроможних скласти основу відповідних лінгвістичних технологій.

1.2. Лексичне значення слова та його лексикографічне представлення

Приклад саме тлумачного словника обрано невипадково. Незважаючи на велике розмаїття типів словників та об'єктів лексикографування, якими в принципі можуть бути будь-які одиниці або відношення мови, центральним об'єктом лексикографування залишається слово, а отже, найпопулярнішими словниками – тлумачні. Тлумачний словник є специфічним системним відображенням реально існуючої мовної дійсності. Мовна система представлена тут у вигляді вилученого із мовлення лексичного ін-

вентарю із відповідними типізованими лінгвістичними характеристиками його одиниць. Дослідження показують, що при певній суб'єктивності в роботі лексикографів відбір слів з їх семантичною, граматичною та стилістичною характеристикою відбувається не випадково, а на основі певних заданих, «нав'язаних» об'єктом принципів, що приводить до системної, закономірної співвіднесеності складу тлумачного словника із загальним складом лексичної системи мови. Зазначені принципи детермінуються об'єктивним улаштуванням мовної дійсності, і лексикографи їх не вигадують, а тією чи іншою мірою враховують, відображають. Таким чином, можна стверджувати, що словник – не хаотичне нагромадження фактів, у ньому певним чином відображаються реально існуючі в мові пласти й сфери словникового складу мови, деякі характерні особливості її системного улаштування. При цьому тлумачний словник демонструє семантичний стан слова як основної одиниці лексикографування – його комплексне представлення, перш за все, як лексичної, граматичної та стилістичної одиниці мовної системи.

Як одиниця лексичної системи слово в тлумачному лексиконі постає як таке, що має лексичне значення (або декілька значень), воно пов'язане різноманітними семантичними відношеннями з іншими лексичними одиницями і з фразеологічним фондом мови, входить у різноманітні лексичні об'єднання та угруповання. Як одиниця граматичної системи слово постає у словнику як морфологічна одиниця, яка відноситься до тієї чи іншої частини мови, має граматичну форму і граматичні значення, а також як одиниця синтаксису, що має свій конструктивний потенціал у сфері сполучуваності слів і побудови речення. Як одиниця стилістичної системи слово характеризується стилістичними позначками, які особливим чином відображають функціональний діапазон та прагматику вживання лексичних одиниць, ступінь їх універсальності або прикріпленості до певних умов спілкування, ступінь їх семантичної специфічності, а також пов'язаний із цим експресивний потенціал їх уживання.

При цьому основним елементом тлумачного словника, підвалиною всієї його структурної надбудови постає лексичне значення слова, яке є одним з найважливіших у мовознавстві та центральним (системотвірним) для тлумачних лексиконів.

Проблеми, пов'язані з категорією значення, незмінно привертають увагу представників зарубіжної, вітчизняної (радянської і пострадянської) лінгвістики, а також представників інших наук: семіотики, логіки й математики (Р. Барт, П. Бріджмен, Р. Карнап, Ч. Морріс, Ч. Піре, Ю. С. Степанов, Г. Фреге, А. Черч та ін.), психології (Л. С. Виготський,

О. Р. Лурія, О. М. Леонтьєв, Е. Рош, С. Л. Рубінштейн та ін.) Подібний інтерес пояснюється тим, що висвітлення питань, пов'язаних зі значенням, дозволяє розкрити онтологічну, соціальну, психологічну й когнітивну природу мови. Інтерес до категорії обумовлюється посиленням позиції антропоцентричної парадигми в гуманітарних дисциплінах, інтеграцією наукового знання, виникненням нових міждисциплінарних наукових напрямів – когнітивної лінгвістики, лінгвоконцептології, лінгвокультурології, нейролінгвістики, психолінгвістики, психосемантики, етнолінгвістики та ін., бурхливим розвитком інформаційно-лінгвістичних технологій, розробкою теорії штучного інтелекту та їхніх застосувань в інформаційних системах та комп'ютерних мережах.

Поняття «лексичне значення» формувалося в класичному мовознавстві під впливом двох фундаментальних понять – «лексика» і «значення».

Поняття «лексика», беручи свій початок у традиційній науці про мову, базується на розмежуванні й протиставленні трьох аспектів мови – звукового складу, лексичних і граматичних засобів. Лексичні засоби – слова і словосполучення, які здавна систематизуються в словниках і називаються словниковим складом (lexicon), поряд із звуковим і граматичним складають основу будь-якої мови і можуть бути названі в цьому смислі її універсальними ознаками. У той же час різниця в репертуарі лексичних одиниць, у своєрідності засобів і типів найменувань, у специфіці семантичного представлення навколишнього світу цими номінативними одиницями понятійних сфер і семантичних категорій складає відмінні ознаки кожної мови. У нашій праці ми виходитимемо з робочого визначення, за яким «слово» є класом мовних одиниць, елементам якого притаманне граматичне та лексичне значення, певна просодична єдність та певна синтаксична функція у реченні. Слово є двосторонньою одиницею, у якій зв'язок значення і звучання (або графічного зображення в системі письма) суспільно, психічно й історично обумовлений і визначає не лише існування, але й розвиток мови.

Якщо поняття «слово» – це основна номінативна й когнітивна одиниця, то його значення формується як «відоме відображення предмета, явища або відношення у свідомості, яке входить у структуру слова як так звана внутрішня його сторона, стосовно до якої звучання слова виступає матеріальною оболонкою, необхідною не тільки для вираження значення і для повідомлення його іншим людям, але й для самого його виникнення, формування, існування та розвитку» [245].

Значення слова обумовлюється усією лексико-семантичною системою мови і є результатом відображення суспільно усвідомленої об'єктивної

дійсності. Лексичне значення формується в умовах конкретних зв'язків і взаємовідношень слів певної мови. На відміну від понять, які є спільними для різних мов, лексичне значення слова завжди національно специфічне як і вся лексика в цілому. Вчені-лінгвісти у своїх працях указують на найважливіший компонент значення – поняття, за допомогою якого встановлюється зв'язок між референтом і словом. Крім поняття, у значення слова можуть входити й інші компоненти: емоційне забарвлення, стилістична характеристика, співвіднесеність з іншими словами мови. На нього нашіраховуються додаткові уявлення й різного роду смислові асоціації. Залежно від того, до якої частини мови належить слово, його лексичне значення пов'язується з певним кодом граматичних значень і може відчувати на собі їхній вплив, оскільки кожна частина мови має свої семантичні особливості.

Аспекти вивчення лексичного значення слова настільки багатоманітні, а дослідницькі підходи такі численні, що вичерпати їх у цьому огляді немає можливості. Загалом слід констатувати, що різні трактування лексичного значення, в якому вчені виділяють та абсолютизують різні його особливості, змушують розглядати природу лексичного значення як складне явище, що включає декілька складових, тобто як специфічну структуру. «Значення слова, – писав В. В. Виноградов, – визначається не тільки відповідністю його тому поняттю, яке виражається за допомогою цього слова (наприклад, рух, розвиток, мова, закон і т. ін.); воно залежить від властивостей тієї частини мови, тієї граматичної категорії, до якої належить слово, від суспільно усвідомлених і усталених контекстів його вживання, від конкретних лексичних зв'язків з іншими словами, обумовлених притаманними певній мові законами поєднання значень слова, від семантичного співвідношення із синонімами й взагалі з близькими за значеннями і відтінками словами, від експресивного і стилістичного забарвлення слова» [129].

Серед одержаної в той чи інший спосіб інформації про слово найбільш широкою і різноаспектною вважається саме та, що міститься у великих глумачних лексиконах, які в словникових дефініціях відображають мовну картину світу, містять відомості про граматичні характеристики лексичних одиниць, про особливості їх функціонування і вирішують головне завдання – з допомогою прийнятого укладачами особливого інструменту опису відобразити основні характеристики семантичної структури слова, системи його значень. Висловлюючи свій погляд на словники та їхню роль, Л. В. Щерба вважав особливо важливим словник сучасної літературної мови народу. В укладанні такого словника він вбачав найважливіше завдання лінгвістів і неодноразово на цьому наголошував [297]. За-

значена настанова досі залишається актуальною, трансформуючись і адаптовуючись до реальності відповідно до когнітивних та технологічних можливостей епохи.

Саме у словниках теоретичні проблеми мовознавчої науки, зокрема проблема опису значення слова, знаходять своє розв'язання, реалізуються і перевіряються результати лінгвістичної теорії, і нерідко словники випереджають теорію практичними вирішеннями ще не до кінця вивчених проблем. Лексикографічні описи самі по собі містять результати лінгвістичного аналізу і завдяки цьому мають велике наукове значення. На важливості словникових праць для розвитку семантичних досліджень наголосувала Н. З. Котлова: «Хоча лексикографічна практика ослаблена тим, що вона не зуміла пов'язати свої рішення з обговоренням теоретичних проблем семантики, їй належить провідна роль в описі лексичної семантики. Успіхи її лексикографічних описів, систематизації й теоретичного осмислення, обумовлюючись прогресом мовознавства, логіки, семіотики, філософії, завжди будуть залежати й від ступеня засвоєння даних лексикографії — і тому, що вся лексикографія представляє результати вивчення великих матеріалів, і тому, що вона відображає всю картину, охоплюючи словниковий склад у цілому, і тому, що лексикографічний опис передбачає одночасне врахування різних властивостей слова. Експеримент такого великого розмаху не може бути повторений поза лексикографією» [177:3]. Саме тому лексикографія перестає бути «мистецтвом укладання словників», як називав її іспанський лексикограф Хуан Касарес, а все більше виявляє себе як особлива наукова дисципліна, зі своїм колом завдань, своїм об'єктом вивчення і своїм призначенням.

Загальновідомо в теорії лексикографії, що внутрішня організація семантики слів подається в словнику відповідно до типу останнього, залежно від його призначення, обсягу, спеціалізації. На протигагу словникам енциклопедичним, які досить широко залучають до своїх реєстрів фахові й спеціальні знання та термінологію, що стосується багатьох галузей науки, техніки, культури і пропонують стисле, але по змозі вичерпне наукове визначення кожного терміна, тлумачні загальномовні словники характеризують слово засобами філологічного тлумачення, розкриваючи суть його семантичної будови. Адже основне завдання будь-якого тлумачного словника полягає в тому, щоб подати якомога точнішу семантичну характеристику лексики даної мови в повному обсязі, який відповідає його призначенню. За словами В. Бланара, «центральною задачею лексикографічного опису є осягнення смислової структури слова на рівні мовної системи» [122].

Проблема представлення семантики реєстрових слів у тлумачній частині загальнономовних словників завжди перебувала в колі уваги вчених-лексикографів. Різноманітні теоретичні й практичні проблеми опису лексичного значення у тлумачній частині словника в європейській, російській та українській лінгвістиці вирішували ряд учених: В. Бланар, Х. Карсарес, Н. П. Денисов, Ю. Д. Апресян, В. Дорошевський, Д. Й. Арбатський, М. Д. Голєв, П. М. Денисов, А. А. Медведєв, Г. А. Уфимцева, Й. А. Стернін, О. П. Євгенєва, Л. І. Кисельовський, Ю. С. Степанов, Н. Ю. Шведова, В. М. Русанівський, Л. С. Паламарчук, О. О. Тараненко, В. А. Широков та ін.

Розвиток лексикографії, яка постійно вдосконалює і збагачує засоби наукового аналізу, сприяє повсякчас глибшому проникненню в семантичну систему мови та ширшому висвітленню процесів взаємодії форми й змісту слова, які не просто передати вербальними засобами. Про це, зокрема, пише Н. Хомський: «Кожному, хто хоч раз намагався дати тлумачення якому-небудь слову, прекрасно відомо, наскільки це важко, — адже щоразу доводиться мати справу з дуже складними й химерними властивостями» [308].

Обсяг і глибина розкриття лексичного значення реєстрової одиниці перебувають у прямій залежності від досягнень лексикографічної науки. Вчені звертають увагу на важливість лексикографічної фіксації семантики слів: «Розуміння принципової важливості методу опису семантичних властивостей слів... лише останнім часом досягло задовільної ясності. У результаті цього намітилося зближення теорії лінгвістичної семантики (лексикології, семасіології) із словниковою практикою, єдиною сферою філологічної праці, у якій проблема семантичного опису вирішується фронтально, на таких лексичних масивах, що семантичні узагальнення, які враховують досвід словникової роботи, стають єдиними переконливими теоретичними висновками» [148:183].

У процесі лексикографічного опрацювання мовленнєва багатоманітність «переплавляється», і на цій основі утворюється науково абстрагована одиниця, відмінна від мовленнєвих слововживань. Лексичне значення стає конкретним фактом об'єктивної дійсності — тлумаченням. Дати слову тлумачення — це дати перифраз, що був би йому «семантично еквівалентним», виступав у формі тексту, складеного відповідно до лексичних і синтаксичних правил мови, із вказівкою на елементи, які дозволяють відрізнити (впізнати, зрозуміти) означуване слово. У кожній дефініції є означувана частина (дефінієндум) і означувальна (дефінієнс). Означувана частина звичайно коротка, вона безпосередньо вказує на те, що піддається

означенню. Означувальна частина містить таку інформацію, яка необхідна для визначення і стає відомою в процесі визначення. Специфічною особливістю дефініції, яка відрізняє її від інших типів висловлювань, є те, що обидві її основні частини знаходяться у певному відношенні еквівалентності. У певних контекстах їх можна поміняти місцями, не порушуючи істинності висловлювання.

Дефініція повинна узгоджуватися як із улаштуванням системи мови, так і з нормою вживання, закріпленою в цій мові. Будучи цілісною одиницею мови, лексичне значення ділиться на частини, які в тлумачних словниках виступають у вигляді рубрик і підрубрик (відтінків). Між основними значеннями і відтінками існують певні відношення. Хоча чітких лінгвістичних розмежувань лексичного значення та його відтінку досі не вироблено, інтуїтивно ясно, що значення можуть мати тонку структуру, яка відображає певні аспекти функціонування дефініційованої одиниці в межах одного семантичного рівня. Відтінки допомагають показати нюанси у взаємовідношеннях лексичних значень, якими характеризується семантична система літературної мови.

Тлумачення як репрезентант лексичного значення у тлумачному словнику є найважливішим компонентом словникової статті, а розробка метамови словника (мови словникового опису), створення точних і несуперечливих словникових тлумачень є найважливішими лексикографічними проблемами: «Укладання дефініцій – вершина лексикографічного мистецтва. У дефініції треба резюмувати зміст слова, а це передбачає тонкий семантичний аналіз, або описати предмет, що вимагає знання реалій та вміння їх розглядати з позицій семіотичної релевантності диференційних та інтегральних ознак [148:191]. Лексикографи різних країн обговорюють теоретично й вирішують практично питання створення універсальної метамови, яка дозволила б, користуючись певним набором семантичних компонентів, створювати тлумачення слів [127:327–328]. Словникове тлумачення і словникова стаття в цілому повинні не тільки в зручній та економній формі дати строгу і по можливості вичерпну інформацію про значення окремого слова, але й показати місце слова в лексичній системі, його зв'язки з близькими за значенням лексичними одиницями, служити основою семантичної взаємодії цієї лексики з іншими у складі висловлювання [109:29].

Основною функцією словникових дефініцій є коректне введення кожного слова реєстру в систему мови з урахуванням його лексичної семантики. Процес лексикографування слова В. Бланар описує в такий спосіб: «Основний хід лексикографічного опису в тлумачних словниках можна

коротко представити у наступному вигляді: за допомогою аналізу лексичних одиниць у тексті (рівень мовлення) ми намагаємося досягнути їхню значимість на рівні мови, а результати комплексного аналізу обробляємо лексикографічно, приймаючи до уваги тип розроблюваного словника» [122]. Означений процес базується на специфіці внутрішньої структури лексичного значення і зумовлює існування певних норм та способів побудови тлумачення, що регулюються правилами, виробленими лексикографічною традицією. Завдання дефініцій у тлумачному словнику сформульовано так: «Тлумачення мають коротко й чітко розкривати значення слова та його основні відтінки і бути по змозі однотипними для слів кожної граматичної категорії» [210:14]. Й. А. Стернін розглядає тлумачню частину, беручи до уваги відбір компонентів значення для неї, формування її структури: «...тлумачення використовує, як правило, такі ознаки значення, які найчастіше актуалізуються» [254]. Тлумачення лексичного значення слова зумовлене звичайно вимогами оптимальності: воно включає такі семантичні компоненти, набір яких є необхідним і достатнім для того, щоб відмежувати одне значення слова від іншого [296].

Центральним завданням лексикографічного опису є осмислення семантичної структури слова на рівні мовної системи: «Загальновідомо, що слово відносить до дійсності, відображає її і виражає свої значення не ізольовано, не у відриві від лексико-семантичної системи цієї конкретної мови, а в нерозривному зв'язку з нею, як її складовий елемент» [129]. При словниковій обробці лексичного значення, тобто при створенні тлумачень, одним із найпродуктивніших є семасіологічний метод опису значень слів шляхом компонентного аналізу, який відіграє провідну роль у лексико-семантичних дослідженнях: «Метод семантичного аналізу слів, який ґрунтується на виділенні компонентів їх значень (компонентний аналіз), вважається одним із найбільш загальноновизнаних способів лексикологічних досліджень» [180:32]. Вважається, що сфера його застосування обмежена більш чи менш чітко організованими лексико-семантичними групами слів, у яких можна виділити набір протиставлень за певними диференційними ознаками. Словникове тлумачення будується на основі виділення мінімальних семантичних компонентів, які дозволяють ідентифікувати об'єкт, включити слово в групу споріднених за смыслом одиниць і чітко визначити відмінності між ними. «Значним досягненням сучасної семасіології є розуміння лексичного значення слова як складного утворення, що складається із ієрархічно організованих семантичних компонентів («атомів» смислу, диференційних семантичних множників, сем і т. д.). Думка про компонентну структуру лексичного значення слова впливає із

розуміння мови як системно-структурного утворення і є надзвичайно плідною та перспективною» [296]. Розкладення значення слова на найпростіші складові походить від ідей Л. Ельмслева і Р. Й. Якобсона та структурної лінгвістики в цілому. Компонентний аналіз, або аналіз за семантичними множниками, розвивався у роботах таких зарубіжних лінгвістів, як А. Греймас, Дж. Лайонз, Ю. Найда, Б. Потт'є, А. Кребер, У. Гуденаф; зокрема російських В. Г. Гака, Й. А. Стерніна, Д. М. Шмельова, Ю. Д. Апресяна, О. М. Селіверстової та ін.; українських Ж. П. Соколовської, Л. А. Лисиченко, Л. О. Кудрявцевої, В. А. Широкова та ін. Цей семасіологічний засіб опису значень слів, що дозволяє дослідити їх семантичну структуру, в сучасному мовознавстві розглядається і як метод дослідження значень слів у лексикології, і як засіб адекватного формулювання тлумачень слів у загальномовному словнику тлумачного типу. Крім того, методика компонентного аналізу передбачає не тільки розкладення значень на складники, а й подальший їхній синтез [179]. При цьому синтез сем у тлумаченні відбувається на якісно новому рівні, в результаті чого виявляються приховані до аналізу сторони об'єкта, визначається його специфіка.

В основі такого опису лежить набір елементарних, далі нерозкладних елементів смислу, які називаються семантичними множниками, або семи. Центральним елементом у семному складі формули тлумачення виступає сема, що є носієм родового значення. Такі семи, на основі яких відбувається ототожнення класу, вчені називають по-різному: архісеми (В. Г. Гак), інтегральні семи (З. Д. Попова, Й. А. Стернін), ідентифікувальні семи (Л. М. Васильєв). Навколо цього центру в певній послідовності розгашовуються необхідні семантичні компоненти, що є носіями видових ознак витлумачуваного слова. Ці більш конкретні семи, які відрізняють індивідуальне значення слова від інших протиставлених йому значень, визначають як диференціальні семи (Л. М. Васильєв), диференційні семи (В. Г. Гак), індивідуалізувальні, діагностувальні семи (Л. В. Арнольд).

У практиці укладання Словника української мови в 20 томах компонентний аналіз застосовується також для встановлення і розмежування таких поширених у мові явищ, як полісемія та омонімія, полісемія та моносемія — на основі критерію наявності чи відсутності загальних семантичних ознак — сем — у зіставлюваних одиницях.

Різні значення слова, а також його відтінки складають його семантичну структуру і служать яскравим прикладом вияву системних зв'язків усередині одного слова. Полісемія, або багатозначність слова, виникає внаслідок того, що мова являє собою систему, обмежену порівняно з безкінечним багатоманіттям реальної дійсності, за словами В. В. Виноградо-

ва: «Мова виявляється змушеною розносити численну множину значень за тими чи іншими рубриками основних понять» [130]. Багатозначність слова — настільки велика і багатопланова проблема, що найрізноманітніші проблеми лексикології та тлумачної лексикографії так чи інакше пов'язані з нею.

Опис полісемії одиниць лексикону (в першу чергу слів) являє собою одне із найбільш складних завдань лексичної семантики. Основні питання наукового опису багатозначності лексичних одиниць пов'язані з визначенням меж цієї категорії. Базові теоретичні завдання у цій галузі можуть бути сформульовані таким чином: 1) розрізнення полісемії та омонімії, тобто встановлення меж, в яких можна говорити про різні значення одного й того самого слова, на відміну від випадків, де маємо справу з різними словами, які збігаються за формою. Полісемія і омонімія є свідченням так званого асиметричного дуалізму мовного знака і відносної свободи означеного і означуваного; 2) розрізнення полісемії та моносемії (тобто встановлення тієї межі, до якої певне вживання слова слід подавати як реалізацію другого значення, на відміну від випадків, коли різницю конкретних уживань слова можна розглядати як контекстно обумовлене варіювання в рамках одного значення).

Визначення меж категорії багатозначності не піддається чіткій операціоналізації. Пошукам критеріїв, які дозволяють розмежувати полісемію та омонімію, з одного боку, полісемію і моносемію — з іншого, посвячено велику кількість досліджень. Однак будь-який із запропонованих критеріїв, узятий окремо, має лише відносний характер. Відносність критеріїв розрізнення полісемії та омонімії, а також відома суб'єктивність у виборі способу словникового опису підтверджується тим, що одні й ті ж слова по-різному трактуються різними словниками: у рамках однієї словникової статті як різні значення — в одних словниках, і як омоніми — в інших. Стосовно полісемії та моносемії можна констатувати, що у різних словниках нерідко зустрічається неоднакове представлення семантичної структури слова.

Зрозуміло, що труднощі у визначенні меж між полісемією та омонімією пояснюються самою організацією мови. Адже реально ми маємо справу не з чітко розмежованими явищами, а з поступовими переходами, тобто з деякою градаційною шкалою, на одному кінці якої знаходяться «класичні» омоніми типу *коса*¹ (заплетене волосся), *коса*² (сільськогосподарське знаряддя), *коса*³ (вузька наливна смуга суходолу в морі, річці), *коса*⁴ (селезінка, діал.), а на другому — тісно пов'язані між собою значення, загальна частина семантики яких має більшу питому вагу, ніж релевантні

семантичні відмінності (пор., наприклад, *поле* у словосполученнях типу *пшеничне поле* і *футбольне поле*).

Що стосується розмежування полісемії та моносемії, то зрозуміло, що в реальному мовленні ми маємо справу з безкінечною множиною різних уживань лексичних одиниць, а не з певними готовими значеннями слів. У якомусь смислі кожне із вживань виявляється унікальним, оскільки слово, що розглядається як одиниця мовлення, тобто використане в конкретній ситуації спілкування, обростає додатковими смислами, принесеними цією ситуацією. Крім того, потрапляючи в певний контекст, воно може підкреслювати, фокусувати деякі ознаки, важливі для цього контексту, і приглушувати, ніби вводити в тінь інші ознаки, потенційно присутні в його значенні. Зведення актуальних, мовленнєвих значень у значення мовні, узуальні і є процесом лексикографічного представлення мовної дійсності.

При цьому застосування семного аналізу в лексикографічній практиці дає змогу подати об'єктивну картину опису системних відношень у мові й розмежувати семантичні структури слів без спільних ознак, встановити подісемію скрізь, де це потрібно для опису лінгвістично значимих властивостей певної лексичної одиниці, постулювати реальні, а не оказіональні значення в лексико-семантичній системі мови.

Дослідження значень слів традиційно здійснюється у рамках окремих об'єднань лексем зі спільною семантикою, що являють собою системи, які по-різному визначаються мовознавцями. Згідно з лінгвістичною теорією поля, крім синтагматичних, функціональних та інших полів, існують парадигматичні поля, до яких відносять найрізноманітніші семантичні класи слів, у тому числі й лексико-семантичні групи. У рамках цих полів формуються окремі семантичні моделі формул тлумачення. Проблема визначення семантики лексичних угруповань отримала ґрунтовну розробку в науковій літературі. Питання однотипного представлення семантики слів окремих парадигматичних груп висвітлюється у працях зарубіжних та українських вчених, зокрема, Й. Тріра, Л. Вейсгербера, К. Ройнінга, Н. Ю. Шведової, Ю. Д. Апресяна, Г. С. Щура, Л. М. Васильєва, О. П. Свєньєвої, Е. В. Кузнецової, Ж. П. Соколовської, Л. А. Лисиченко, О. О. Таранька та ін. До парадигматичних об'єднань учені відносять різноманітні класи слів: тематичні, лексико-семантичні, синонімічні, антонімічні та інші групи.

Важливою проблемою у сучасній лексикографічній науці є обробка представлення семантики окремих семантичних об'єднань слів. Додатність розроблення однотипних тлумачень слів вбачала Н. Ю. Шведова для окремо взятої частини мови: «Для того щоб уникнути різнобою у

словнику, лексикограф повинен працювати... з певними лексичними групами в межах окремих частин мови. Тільки на цьому шляху можна досягти єдності у розмежуванні значень, однотипності тлумачень» [268]. Автор наголошує на тому, що у результаті словник стане науковим твором, який спиратиметься на спеціально розроблені «еталони опису слів, об'єднані формально і семантично» [268]. На необхідності уніфікованої побудови дефініцій наголошувала також О. П. Євгенєва, яка визнавала шлях створення моделей дефініцій саме для «тематичних» груп лексики, з огляду на те, що такі спроби спираються на узагальнювальну і класифікувальну силу людської свідомості [153]. Дослідниця лексико-семантичної системи Е. В. Кузнецова серед інших парадигматичних об'єднань лексики виділяє саме лексико-семантичні групи: «Найважливішим з погляду лексичної системи типом класів слів є лексико-семантичні групи. Вони об'єднують у собі слова однієї частини мови, в яких, крім загальних граматичних сем, є як мінімум ще одна загальна сема – категоріально-лексична (архісема, класема). Категоріальна сема передбачає, «задає» не будь-які, а якісь певні аспекти свого уточнення. У рамках цих аспектів формуються типові диференційні семи. У зв'язку з цим у кожній окремій лексико-семантичній групі набір диференційних сем виявляється специфічним [180].

Питання необхідності вивчення словникової інтерпретації лексичних значень слів у межах певних угруповань порушувалося також і в українській лексикографії. У праці В. А. Широкова акцентується увага на тому, що моделювання лексичної семантики полягає у виявленні мовних засобів – слів і зв'язків між ними, за допомогою яких можлива декомпресія лексичного значення реєстрових одиниць словника на основі формул тлумачення, та у встановленні співвідносних за тими чи іншими виокремленими параметрами угруповань [275:176]. Отже, лексичний склад мови в тлумачному словнику одержує системний опис, що виявляє і репрезентує зв'язки та відношення слова з іншими словами у межах лексико-семантичних груп.

До складу парадигматичних об'єднань відносяться лексико-семантичні поля – групи слів зі спільною гіперсемою, яка задає гіперонімічне відношення, що є визначальним для слів цього поля. Різні лексико-семантичні поля можуть мати непорожній перетин; багатозначні слова своїми значеннями можуть входити до різних полів. У свою чергу, лексико-семантичні поля розпадаються на семантично дрібніші об'єднання – лексико-семантичні групи, у межах яких виділяють об'єднання, побудовані на відношеннях синонімії, антонімії, конверсії та гіпо-гіперонімії. У лексико-семантичну групу об'єднуються слова однієї частини мови на ос-

нові єдиної, узагальненої семантичної ознаки, що міститься у значеннях цих слів. Наявність такої спільної ознаки виражається у семантичній співвіднесеності всіх елементів лексико-семантичної групи, що виступає її суттєвою ознакою. Це має пряме відношення до поняття лексикографічного типу: до однієї групи лексем, об'єднаних хоча б однією якоюсь спільною властивістю, застосовуються одні й ті самі правила опису в словнику [110]. Таким чином простежується типове представлення елементів значення слова у тлумаченнях слів однієї лексико-семантичної групи, а сама формула тлумачення виступає певним чином стандартизованою, уніфікованою структурою, в межах якої реалізується однотипне функціонування семантичних компонентів. У зв'язку з цим актуальними є поділ складу словника на лексико-семантичні об'єднання та розробка у їх складі уніфікованих формул тлумачення. Такий напрямок досліджень є актуальним і конструктивним для формування наукової парадигми лексико-семантичних відношень мови, їхньої класифікації та комбінації, подання у вигляді цілісної картини: «Що стосується сучасної семантики, то основним предметом її уваги є семантичне представлення: замість того, щоб говорити про значення (і зміни значення), вона прагне моделювати їх і представити у вигляді експліцитних формул» [126].

Продовжуючи традицію укладання великих словників тлумачного типу, Словник української мови у 20 томах, суттєво модифікований, уточнений і доповнений, має на меті представити сучасну українську мову в усій багатоманітності та повноті лексичного складу, відобразити сучасні уявлення українців про світ, передати нащадкам духовні надбання нації. За твердженням Н. Ю. Шведової, «словник розгортає перед своїми читачами картину світу – так, як вона існує у свідомості носіїв цієї мови» [268].

1.3. Концепція станів мовних одиниць

Першим питанням, що виникає при моделюванні мови, є питання про об'єкти моделювання, а саме: що являють собою об'єкти мови й що, власне, ми збираємося моделювати. Враховуючи вищевикладене, за вихідне візьмемо положення про те, що власними (елементарними) об'єктами концептуального представлення в лінгвістиці мають виступати не безпосередньо одиниці мови (одиниці фонетичного, морфологічного, лексичного та інших рівнів, які ми інтерпретуємо як онтологічні об'єкти), а певні «проміжні» стосовно мови об'єкти, феноменологічними корелятами яких є психофізичні стани й процеси, що мають місце в мовнорозумовому апараті людини, так що усна й писемна її форми служать елементами інфра-

структури мовного процесу. Опису й моделюванню підлягають як, власне, згадані психофізичні стани й процеси, так і їхня інфраструктура.

Застосування поняття стану мовної одиниці одержало розвиток у роботах [275; 283], згідно з якими будь-яке слово (загалом кажучи – будь-яка одиниця мови) у контексті або в мовному потоці перебуває в певному семантичному стані¹. Для одиниць лексичного рівня цей стан являє собою певну суму ознак граматичної та лексичної семантики і надає шлях для узагальнення понять граматичного й лексичного значення. Власне, сам процес розуміння мови з цієї позиції виглядає як редукція апріорного розподілу лексем за сумарними ознаками граматичної та лексичної семантики, які властиві суб'єктивному лексикону реципієнта, до певного одного граматичного й лексичного значення, характерного саме для того контексту, який перебуває у полі уваги реципієнта і підлягає в даний момент процесу його індивідуальної мовної обробки.

Незважаючи на те, що й дотепер не створено формального апарату для визначення системи всіх можливих семантичних станів будь-якої лексеми, природно вважати, що досить адекватною моделлю даної системи може служити комплекс граматичних і лексичних значень – так, як він представлений, скажімо, у великому тлумачному словнику тієї або іншої мови.

Таким чином, при розгляді формальних аспектів семантики будемо виходити з існування відповідності між мовною одиницею та її станом:

$$s : X \rightarrow s(X), \quad (1.1)$$

де X – деяка одиниця мови; s – відповідність між X і $s(X)$ – формальним об'єктом, що представляє семантичний стан одиниці X , який має своїми детермінантами елементи засобів матеріального вираження семантики. Для будь-якої одиниці X семантичні стани утворюють певну множину – вважатимемо її скінченною, але не обмеженою, і позначимо символом $\{s(X)\}$. Клас певних одиниць мови L позначимо символом $W(L)$ або просто W , якщо йдеться лише про одну конкретну мову; належність X до W позначимо: $X \in W$; множину всіх семантичних станів для всіх $X \in W$ позначимо через $S = \{s(X); X \in W\}$.

Припустимо, що існує оператор F , дія якого визначена на множині семантичних станів S і який будемо інтерпретувати як оператор значення певної семантичної категорії (ними можуть бути, до речі, й категорії граматичної семантики – якщо, наприклад, F є оператором частини мови, то

¹ Терміни *стан* і *семантичний стан* надалі використовуватимуться як синоніми там, де це не призводить до плутанини.

його значеннями є: f_1 – іменник, f_2 – дієслово, f_3 – прикметник і т. д.). Це означає, що F являє собою певний інтелектуальний механізм, що, аналізуючи семантичний стан $s(X)$, у якому перебуває одиниця X , ідентифікує певне значення конкретної семантичної категорії, що відповідає саме даному семантичному стану. Зрозуміло, що встановлення системи операторів типу F повинно ґрунтуватися на певній семантичній теорії, формалізованій належним чином.

Математичною мовою дію оператора F можна виразити в такий спосіб:

$$F s_i(X) = f_i s_i(X), \quad (1.2)$$

де f_i – деяке значення семантичної категорії; функції $s_i(X)$, $i = 1, 2, \dots$ представляють семантичні стани одиниці X , що маркують її належність до значення f_i категорії F . Значення f_i , $i = 1, 2, \dots$ будемо називати власними значеннями оператора F , що відповідають семантичним станам $s_i(X)$, $i = 1, 2, \dots$. Множину всіх семантичних станів, що відповідають власному значенню f_i , називатимемо множиною часткових семантичних станів і позначимо символом $S(f_i)$:

$$S(f_i) = \{s : Fs = f_i s_i\}. \quad (1.3)$$

За визначенням множина $S(f_i)$ складатиметься тільки з тих семантичних станів, які характеризуються певним значенням семантичної категорії F , а саме – значенням f_i .

Рівняння (1.2) у тих випадках, коли множина, що пробігає індекс i , складатиметься більше ніж із одного елемента, представляє формальне вираження явища семантичної багатозначності. Для одиниць лексичного рівня це полісемія й омонімія (як лексична, так і граматична). Що, зокрема, у випадку омонімії відображає ситуацію, коли словоформа X в одному контексті може перебувати, наприклад, у стані іменника, а в іншому – в стані дієслова. Такий різновид омонімії звичайно називається міжчастиномовною омонімією.

Можливі випадки, коли в рівнянні (1.2) певному значенню f_i відповідає не один семантичний стан $s_i(X)$, а кілька: $s_j(X)$, $j = 1, 2, \dots$. Такі стани будемо називати виродженими; про виродженість сигналізує наявність верхнього індексу біля символу семантичного стану $s_j(X)$. Наприклад, в українській мові словоформа «МАТИ» зі значенням категорії *частина мови* «іменник» має два граматичних стани $g(\text{МАТИ})$:

Гіменник *жін. рід, однина, наз. відм.*

де слово «МАТИ» має лексичне значення «жінка стосовно до дитини, яку вона народила», і:

Семантик чол. рід, множина, наз. відм.,

де слово «МАТИ» має лексичне значення «спортивні матраци». Цей приклад ілюструє явище міжчастиномовної омонімії.

Кількість власних семантичних станів одиниці X , що відповідають певному значенню категорії F , будемо називати кратністю виродження цього семантичного стану. Так, у розглянутому прикладі стан $g(\text{МАТИ})$ має кратність вираження 2:

Семантик жін. рід, однина, наз. відм.

та

Семантик чол. рід, множина, наз. відм.

Семантичні стани, у яких оператор F має певне значення, будемо називати чистими. Однак, а ргоіг не можна забороняти існування семантичних станів, для яких оператор F не набуває одного певного значення, але може характеризуватися, наприклад, двома. Формально таку ситуацію можна зобразити за допомогою співвідношення:

$$F(X) = f_1 \alpha_1(s_1) s_1(X) + f_2 \alpha_2(s_2) s_2(X), \quad (1.4)$$

де семантичний стан s мовної одиниці X при дії на нього оператора F розщеплюється на два, а саме: $s_1(X)$ і $s_2(X)$, де $s_1(X)$ відповідає значенню семантичної категорії f_1 , а $s_2(X)$ — значенню семантичної категорії f_2 ; лінгвістичний зміст функцій $\alpha_1(s_1)$ і $\alpha_2(s_2)$ буде з'ясований нижче.

Стани, для яких оператор F , діючи на функцію семантичного стану, зображується комбінацією певного числа чистих семантичних станів, що відповідають різним власним значенням цього оператора, будемо називати змішаними семантичними станами.

Таким чином, рівняння (1.4) визначає семантичний стан, у якому перебуває одиниця X і яке являє собою своєрідне явище суперпозиції (суміші) чистих семантичних станів $s_1(X)$ і $s_2(X)$, що відповідають власним значенням f_1 і f_2 відповідно. Лінгвістична інтерпретація рівняння (1.4) полягає в тому, що мовна одиниця X має семантичні ознаки одночасно й f_1 і f_2 . Відповідні показники — ідентифікатори відзначеної належності знаходяться у вираженнях для семантичних станів $s_1(X)$ і $s_2(X)$.

Така ситуація досить типова в мові. Наприклад, в українському й російському дієприкметниках з'єднані властивості дієслова й прикметника. Розглянемо, наприклад, російські прислівникові лексеми *ведущий* і *ведомый*. Вони відмінюються за словозмінною парадигмою прикметника (по

ність відмінків у чоловічому, жіночому й середньому роді та у множині), маючи одночасно у своїй структурі дієслівну морфологічну ознаку активності / пасивності – вона набуває тут матеріального вираження у вигляді суфіксів *-ущ* і *-ом*, відповідно. Відзначена морфологічна ознака не виглядає сильною з погляду належності до дієслова, де вона не набула статусу словозмінної; крім того, вона характерна для повної словозмінної парадигми, а не тільки для окремих її членів – це дає підставу для виділення класу слів з такими властивостями в самостійну частину мови – *причастие*. В українській мові *причастие* є *дієприкметник*, і в цій назві знайшли своє відбиття обидві ознаки – дієслова й прикметника².

Підкреслимо принципову відмінність між описаною ситуацією та звичайно омонімією. Явище омонімії також можна описати рівнянням типу (1.4): $Fs(X) = f_1\alpha_1(s_1)s_1(X) + f_2\alpha_2(s_2)s_2(X) + \dots$, де різні члени правої частини відповідають різним омонімічним станам слова X – якщо йому властивий той або інший різновид омонімії. Але в процесі мовної обробки, коли відбувається зняття омонімії, права частина цього рівняння редукується до одного члена, що, власне, і представляє чистий граматичний стан аналізованого слова X у конкретному контексті. Зовсім інша ситуація у випадку, коли X перебуває в змішаному стані, – у цьому випадку ніяка мовна обробка не здатна зменшити число членів рівняння (1.4), що й представляє її кінцевий результат. Таке становище сигналізує про існування одиниць, які в контексті функціонують у змішаних семантичних станах.

Висновок про існування одиниць, для яких мова допускає функціонування тільки в змішаних семантичних станах, виглядає досить цікавим свідченням неможливості повної формалізації мовної системи, точніше – свідченням меж цієї формалізації. Це, зокрема, означає, що навіть найбільш формальна область теорії мови, якою є граматики, має риси нечіткості, що входить у певне протиріччя із звичкою до граматичного детермінізму і стимулює напрацювання спеціальної мови опису багатозначних граматичних ситуацій.

Концептуальну базу такої мови, на нашу думку, надає теорія нечітких множин Л. Заде [160]. Продемонструємо, в який спосіб це досягається.

Звернімося до визначення S семантичних станів і множин $S(f_i) = \{s: Fs = f_i s_i\}$ часткових семантичних станів. Якби існувала принципова можливість редукції у процесі мовної обробки будь-якого семантичного стану

² Цікаві приклади суперпозиції граматичних станів у турецькій мові наведено в роботі [293], де іменник має певні категоріальні та матеріальні ознаки дієслова.

до чистого, то множину S можна було б представити у вигляді об'єднання множин $S(f_i)$, які між собою не перетинаються, тобто справедливою була б формула:

$$S = \bigcap_i S(f_i); S(f_i) \cap S(f_j) = \emptyset \text{ при } i \neq j. \quad (1.5)$$

Ситуація зовсім інша, якщо ми врахуємо описану можливість існування одиниць, що одночасно характеризуються, наприклад, двома або більше значеннями певної семантичної категорії. Семантичний стан $s(X)$ такої одиниці X уже не буде належати лише одній із підмножин $S(f_i)$, але одночасно двом, а можливо, навіть і більше.

Формальний механізм опису подібних явищ здійснюється в такий спосіб. Визначимо на множині $S = \bigcap_i S(f_i)$ структуру нечіткої в смислі Заде множини. Для цього на кожній з підмножин $S(f_i)$ визначимо функцію належності $\alpha_i(s)$, що для кожного $s \in S(f_i)$ набуває певного числового значення з інтервалу $[0, 1]$:

$$\alpha_i(s) \in [0, 1]. \quad (1.6)$$

При цьому будемо вважати, що якщо $\alpha_i(s) = 1$, то стан s є чистим, а якщо $\alpha_i(s) < 1$, що відповідає s , є компонентом змішаного стану, другий компонент якого s' (для двокомпонентних станів) належить певній підмножині $S(f_j)$, $i \neq j$, зі значенням функції належності $\alpha_j(s')$, також меншим 1, але таким, щоб виконувалася умова:

$$\alpha_i(s) + \alpha_j(s') = 1. \quad (1.7)$$

Тоді кожна з підмножин $S(f_i)$ перетворюється на нечітку множину із функцією належності $\alpha_i(s)$:

$$S(f_i) \rightarrow \{S(f_i), \alpha_i(s), s \in S(f_i)\}. \quad (1.8)$$

Елементами певної нечіткої множини служать пари $(s(f_i), \alpha_i(s))$. Одночасно структура нечіткої множини індукується і на всій множині S як об'єднанні нечітких множин (1.8) при виконанні умови (1.7).

Функція належності $\alpha_i(s)$ при цьому одержує інтерпретацію як міра отримання семантичним станом s властивостей значення f_i категорії F . Максимальне значення, рівне 1, функція належності набуває на чистих станах, що характеризує слова з однозначно певним показником належності до певної семантичної категорії.

Для слів зі змішаними семантичними станами виду:

$$s(X) = \alpha_1(s_1)s_1(X) + \alpha_2(s_2)s_2(X) \quad (1.9)$$

величина $\alpha_1(s_1)$ демонструє ступінь, з якого мовний об'єкт X проявляє властивості значення f_1 , а $\alpha_2(s_2)$ відповідно, ступінь, з якого X проявляє

властивості значення f_2 . Умова (1.7), а саме: $\alpha_1(s_1) + \alpha_2(s_2) = 1$, забезпечує повноту семантичних властивостей розглянутого об'єкта й замкненість семантичного опису.

1.4. Формалізований підхід до визначення семантичних станів слів

Поряд із теоретичними положеннями щодо природи граматичного і лексичного значення та їх рефлексивними визначеннями не менш важливим для лексикографії є операціональний аспект цих понять, який вимагає формалізації дій, пов'язаних із процесом абстрагування граматичних та лексичних значень. А саме, постає питання: в який спосіб (способи) можна встановити різні значення слова, за допомогою яких засобів їх можна розрізнити, як встановлюються відношення між ними, зокрема, відношення близькості або, скажімо, протилежності, як визначаються їх семантичні детермінанти, якими, власне, ці детермінанти можуть бути, і т. ін.

Очевидно, що значення слова розкриваються тільки в контексті. Отже, теоретично для експлікації значень конкретного слова необхідно зібрати всі — у певному сенсі — його контексти, де воно функціонує, розподілити їх за однорідними у певному (семантичному) відношенні групами, кожна з яких і є репрезентантом певного лексичного значення. Далі, вивчаючи ці групи контекстів, лексикограф виводить із кожної такої групи окреме лексичне значення аналізованої лексеми і кваліфікує відповідні граматичні значення.

Зрозуміло, що така схема є ідеалізацією, адже сума контекстів є нескінченною, а на практиці лексикограф завжди має справу з обмеженим (часто дуже обмеженим) колом контекстів, з якого він, спираючись на своє знання системи мови, лінгвістичний досвід та інтуїцію, вибудовує систему лексичних значень. Зрозуміло, що при цьому важко позбавитися суб'єктивізму, притаманного будь-якому, навіть дуже об'єктивно налаштованому, дослідникові.

Постає ще декілька питань. Чи будь-якому контексту $k(x)$ слова x відповідає певне лексичне значення? Чи не може виникнути ситуація, коли одному й тому самому контекстові мусимо приписати два лексичних значення або й більше, чи взагалі приписати значення, які внаслідок десемантизації або якихось інших семантичних процесів не можна трактувати як лексичні? Відповіді на ці питання, типові для когнітивних ситуацій, характерних для формалізму семантичних станів, можна одержати лише

шляхом ретельного аналізу досить великих, репрезентативних колекцій контекстів.

Такими зібраннями, що слугують фактологічною базою визначення лексичної семантики, у лексикографії традиційно використовувалися лексичні картотеки, в яких суцільно або вибірково розписувалися на лексичні картки певні твори – художні, публіцистичні, наукові, офіційно-ділові, фольклорні й т. ін. Вони розташовувалися в алфавітному порядку слів – лінгвістичних центрів відповідних контекстів і вивчалися лексикографами при встановленні та описі ними лексичної семантики. Як було зазначено вище, саме такою є шестимільйонна картотека, створена в академічному Інституті мовознавства ім. О. О. Потебні, масиви якої було використано при укладанні СУМ-11. У наш час завдання щодо зібрання та експлікації контекстів покладається на лінгвістичні корпуси – значні за обсягами електронні бібліотеки текстів, які розмічаються (маркуються) граматичними параметрами всіх наявних у цих текстах слів, що складають спеціальний пошуковий індекс, використання якого й надає можливість практично миттєвого одержання повного набору контекстів будь-якого об'єкта слова за всіма джерелами, представленими в корпусі. Великі лінгвістичні корпуси сьогодні мають обсяги сотні мільйонів слововживань – комп'ютерних аналогів традиційних лексичних карток. Докладний опис конструкції, принципів побудови лінгвістичних корпусів та застосування корпусу при створенні СУМ-20 викладено в монографії [283] та у шостому розділі цієї книги.

На основі вивчення в той чи інший спосіб одержаних сукупностей контекстів здійснюється формування комплексів лексичних значень слів, що є об'єктами лексикографування.

Викладене підсумовано у вигляді певної схеми або послідовності етапів.

На першому етапі формується множина $K(x)$ контекстів мовної одиниці (лексеми³) x , де x пробігає клас слів певної мови. У такий спосіб одержується «множина множин» контекстів.

На другому етапі множина $K(x)$ для кожної лексики x розподіляється за сукупностями неперетинних підмножин:

³ Тут лексемою (на відміну від визначення лексики, прийнятого Московською семантичною школою) називається слово зі знятою омонімією разом із сукупністю всіх його словозмінних форм, яким притаманний певний спільний комплекс лексичних значень.

$$K(x) \supseteq \bigcup_{i=1}^{N(x)} K_i(x); K_i(x) \cap K_j(x) = \emptyset \text{ при } i \neq j, \quad (1.10)$$

де $N(x)$ – ціле число, яке ототожнюється із кратністю полісемії лексеми x . Здійснення розбиття $K(x)$ на окремі підмножини $K_i(x)$, $i=1, 2, \dots, N(x)$, є неформальним інтелектуальним процесом, оскільки саме у ньому відбувається зіставлення множині $K_i(x)$ певного лексичного значення. Іншими словами, ми припускаємо, що між лексичним значенням $l_i(x)$ та множиною контекстів $K_i(x)$ існує певна взаємнооднозначна відповідність:

$$l_i(x) \leftrightarrow K_i(x), \quad (1.11)$$

яку встановлює і фіксує лексикограф. При цьому для простоти викладу вважаємо, що множини $K_i(x)$ сформовано в такий спосіб, що в кожному контексті $k_i'(x) \in K_i(x)$ лексема x , хоча й може зустрічатися більше одного разу, але в одному й тільки одному лексичному значенні.

Зробимо деякі зауваження, оскільки насправді тут ми маємо більш складний процес. На практиці не всі контексти лексичної одиниці x беруться до розгляду при визначенні лексичних значень. Розглядаються й аналізуються лише ті контексти, в яких значення відповідної одиниці є «соціалізованим». Тобто таких контекстів є досить багато і вживання розглядуваної одиниці в них є суспільно усталеним, яке можна ввести й закріпити в конвенціональному модусі. Занадто індивідуалізовані, «оказіональні» контексти, зазвичай, кваліфікуються як виразники «сміслів». У такий спосіб екстралінгвістичний зміст лексичної одиниці набуває двоїстасного характеру, а саме, як «значення» та «смісл»⁴.

Таким чином, слід констатувати, що у формулі (1.10) підмножини $K_i(x)$, $i=1, 2, \dots, N(x)$, взагалі кажучи, не повністю покривають всю множини контекстів $K(x)$. Разом і поряд з $K_i(x)$ у мовленні можуть існувати й функціонувати певні «виняткові» контексти, які маніфестують змісти, що не зводяться до «стандартних» лексичних значень, приписуваних контекстам $K_i(x)$. Строго кажучи, формулу (1.10) варто було б переписати у такому вигляді:

⁴ Порівняємо, для прикладу, сказане із визначенням значення та смислу, яке дає І. М. Кобозєва [171:13]: «Значение X-а – это информация, связываемая с X-ом конвенционально, т.е. согласно общепринятым правилам использования X-а в качестве средства передачи информации».

$$K(x) = \bigcup_{i=1}^{N(x)} K_i(x) + Q(x), \quad (1.12)$$

де $Q(x)$ — множина «нестандартних» контекстів, у яких маніфестовано змісти, що не зводяться до конвенціональних лексичних значень, приписуваних контекстам $K(x)$. Розбиття $K(x)$ на дві групи, подане у формулі (1.12), а радше існування (хоча б і потенційне) множини «смислів» $Q(x)$, є свідченням семантичної незамкненості мовної системи та практично необмежених виражальних можливостей мови. Як наслідок викладеного, констатуємо, що контексти $Q(x)$ є «несловниковими» і стандартно не лексикографуються.

Зауважимо, що формулювання лексичного значення $l(x)$ (або побудова його словникової дефініції) відбувається вже на наступному — третьому етапі. Саме тут на підставі аналізу членів множини $K(x)$ здійснюється категоризація елементів словникової дефініції, їх формалізація і облаштування дефініції у вигляді, який відповідає концепції конкретного словника й лексикографічній традиції.

Сукупність зазначених етапів в основному і складає процес лексикографування.

Як було сказано, перший етап полягає у формуванні множини контекстів $K(x)$ лексеми x , коли x пробігає клас слів певної — у нашому випадку української — мови. Позначимо цей клас символом W^U , тобто $x \in W^U$. Вважаємо, що клас W^U містить всі українські слова. Їх число вважатимемо скінченим, але необмеженим. Скінченність W^U є цілком зрозумілою властивістю, адже у протилежному разі ми мусили б констатувати наявність серед членів W^U таких елементів, які мають нескінченну довжину, що неможливо. У той же час, W^U слід вважати необмеженою сукупністю, адже у будь-який момент її можна розширити шляхом додавання якогось нового елемента.

Для формування множини $K(x)$ необхідно мати певну множину текстів (усних та / або писемних), які були б проіндексовані елементами $x \in W^U$. Така індексація забезпечує прямий доступ до всіх контекстів обраної лексеми x . Кожен конкретний контекст $k(x)$, який належить до $K(x)$: $k(x) \in K(x)$, є текстом певної довжини, що містить слово x . Довжина $k(x)$ мусить бути достатньою для того, щоб визначити значення слова x . Як правило, ця довжина обмежується одним реченням, проте трапляються випадки, коли одного речення недостатньо. У лексичних картотеках визначення довжини контексту становило досить непросту проблему. В су-

часних лінгвістичних корпусах довжина контекстів не є сталою величиною і може змінюватися за бажанням користувача.

З формули (1.11), зокрема, випливає, що всі контексти з множини $K_i(x)$ є еквівалентними: $k_i^r(x) \sim k_i^p(x)$, $k_i^r(x) \in K_i(x)$; $k_i^p(x) \in K_i(x)$, по відношенню належності до $l_i(x)$. Однак в різних контекстах $k_i^r(x)$, $r = 1, 2, \dots$, лексема x може набувати різних граматичних значень. Тому, використовуючи формалізм теорії семантичних станів, можна подати диференціальний семантичний стан $s_i^r(x)$ лексеми x в контексті $k_i^r(x)$ у такому вигляді:

$$s_i^r(x) = g_i^r(x)l_i(x), \quad (1.13)$$

де символом $g_i^r(x)$ позначено граматичний стан (формальну репрезентацію граматичного значення) лексеми x в контексті $k_i^r(x)$, а через $l_i(x)$ – лексичний стан (формальну репрезентацію лексичного значення) лексеми x у цьому ж контексті. Очевидно, що зміна індексу r , тобто варіація граматичного значення лексеми x , не впливає на її лексичне значення. Це означає, що можна визначити семантичний стан лексеми x у вигляді зваженої суми:

$$s_i(x) = \left[\sum_r \alpha_i^r(x) g_i^r(x) \right] l_i(x), \quad (1.14)$$

де індекс r маркує різні граматичні значення лексеми x , а ваги $\alpha_i^r(x)$ вибрано так, що $\sum_r \alpha_i^r(x) = 1$; $\alpha_i^r(x) \geq 0$.

Співмножник $\sum_r \alpha_i^r(x) g_i^r(x)$ у формулі (1.14) відповідає за внесок окремих граматичних значень у семантичний стан лексеми x . Зауважимо, що на сьогодні нам невідомі праці, де б проводилося диференціальне дослідження внеску окремих граматичних значень у загальну семантику лексем.

Формулу (1.13) можна узагальнити на випадок, коли необхідно врахувати взаємодію граматичної та лексичної семантики:

$$s_i^r(x) = g_i^r(x) I(g; l; x) l_i(x), \quad (1.15)$$

де член $I(g; l; x)$ відповідає за взаємодію між граматичною та лексичною семантикою лексеми x у стані $s_i^r(x)$.

Розглянемо окремо випадок, коли різні лексеми характеризуються єдиним комплексом лексичних значень. Причому тут ідеться про такі лексеми, які не є абсолютними синонімами, а варіюють за певною граматичною ознакою, що не може кваліфікуватися як словозмінна. Приклади такого типу досить поширені, зокрема, у системі українського дієслова [225]. Ілюстрацією може бути словникова стаття із реєстровою одиницею НАЛАШТОВУВАТИ із СУМ-11:

НАЛАШТОВУВАТИ, ую, уеш, *недок.*, **НАЛАШТУВАТИ**, ую, уеш, *док.*, *що*. 1. Готувати що-небудь або приводити до стану готовності. *Сакен почав налаштувати парус з таким розрахунком, щоб поманити ще турків* (С. Добровольський); — *Не знаєте, чи хазяїн уже устали? — спитався він.* — *Казали, сьогодні кудись їхати, так щоб віз налаштувати* (Панас Мирний); *Уляна висипала з пелени квасоллю і вже хотіла йти до хати, щоб налаштувати їжу кабанцеві* (Григорій Тютюнник); // *кого*. Настроювати кого-небудь на певний лад, настрої і т. ін. *Галина ж, ідучи сюди, налаштувала себе на звичайність, навіть мимохідність свого завітання* (М. Олійник).

2. *що, чого*. Робити що-небудь придатним для роботи, користування і т. ін. *Під час вечері Мишуня налаштував радіоприймача* (Ю. Яновський); — *Братчики, до гармат! — закричав Гулик. З ним було до десятка старих гарматців, які одразу ж кинулись до борожих єдинорогів і почали налаштувати їх до стрільби* (С. Добровольський).

Ліва частина цієї словникової статті **НАЛАШТОВУВАТИ**, ую, уеш, *недок.*, **НАЛАШТУВАТИ**, ую, уеш, *док.* представлена видовою парою «налаштувати» (дієслово недоконаного виду) і «налаштувати» (дієслово доконаного виду). Зазначені дієслова, очевидно, належать до різних дієслівних словозмінних класів [270] і мають різні словозмінні парадигми [290]. Отже, граматична семантика вмотивовує належність цих дієслів до різних лексем. Проте вони характеризуються спільною системою лексичних значень, що видно з лексикографічної розробки наведеної словникової статті. Таким чином, ці дієслова характеризуються одним повним семантичним станом:

$$x \text{ (налаштувати)} = \sum_i \beta_i s_i \text{ (налаштувати)}, \sum_i \beta_i = 1, \beta_i \geq 0, \quad (1.16)$$

де s_i (налаштувати) є семантичним станом, визначеним згідно з формулою (1.14), в якій, однак, слід здійснити певне уточнення, що стосується поняття «лексема». А саме це поняття варто узагальнити на сукупність усіх варіантів форм слова x , які мають спільний комплекс лексичних значень. Отже, поняття семантичного стану є більш загальним, аніж поняття граматичного та лексичного значення. В українській мові зазначена вище варіативність може здійснюватися за різними граматичними морфологічними параметрами. Наприклад:

Префіксальна варіація:

а) вбїк ~ убїк, вбїлений ~ убїлений, вбогїсть ~ убогїсть, вболївати ~ уболївати, (фонетична префіксальна варіація);

б) згїянутися ~ їзгїянутися ~ зогїянутися; згарячу ~ їзгарячу; ззадү ~ їззадү; зобганїй ~ їзобганїй; відбївати ~ одбївати (структурно-фонетична префіксальна варіація).

Коренева варіація:

кмітити ~ кметити; кмітувати ~ кметувати.

Суфіксальна варіація:

а) устрінуги ~ устрити (не веде до зміни словозмінного класу і виду дієслова);

б) **БУЛЬКОТАТИ** (очу́, б́чеш) ~ **БУЛЬКОТИТИ** (очу́, оти́ш) – (веде до зміни словозмінного класу, але не веде до зміни виду дієслова);

в) **ВИГОВО́РЮВАТИ** (недок.) ~ **ВИГОВОРИТИ** (док.) – (веде до зміни виду дієслова).

Можливе також і комбінування різних типів варіативності. Отже, варіативність, що має граматичне походження, може бути типізована і формально врахована у члені $\alpha_i'(x)g_i'(x)$ формули (1.14) через введення спеціального параметра l , за допомогою якого здійснюється маркування різних типів варіативності лексеми x , що зберігають комплекс її лексичних значень:

$$\sum_r \sum_l \alpha_i''(x)g_i''(x^l). \quad (1.17)$$

У третьому розділі монографії буде побудовано формалізовану теорію типів варіативності, які можуть зберігати лексичні значення для українського дієслова та частково – для іменника. Причому для дієслова можна буде точно визначити, перерахувати та специфікувати всі типи варіативності лексичних одиниць, для яких можливе збереження повного комплексу лексичних значень. Тобто буде виведено необхідні умови для забезпечення зазначеного збереження.

У результаті формула для представлення семантичного стану лексеми з урахуванням описаної варіативності, що не веде до зміни повного лексичного значення, набуває вигляду:

$$s(x) = \sum_l \beta_l \left[\sum_r \sum_i \alpha_i''(x)g_i''(x^l) I(g; l; x^l) \right] I_l(x), \quad (1.18)$$

де коефіцієнти β_l можна інтерпретувати як відносні ваги (середні частоти), з якими різні лексичні значення $I_l(x)$, $l = 1, 2, \dots$, зустрічаються у множині $K(x)$.

Згідно з викладом попереднього підрозділу, формалізований опис лексичних значень буде в основному здійснений за Ю. Д. Апресяном [108], який, використовуючи семіотичний підхід Ч. Морріса [311] та його лінгвістичну інтерпретацію І. А. Мельчуком [193], запропонував при визначенні лексичних значень урахувати семантичні, синтактичні та прагматичні аспекти. У нашому формалізмі через апарат станів ця пропозиція може бути представлена у такий спосіб:

$$I_i(x) = \sigma_i(x)Z_i(x) + \tau_i(x)U_i(x) + \pi_i(x)V_i(x). \quad (1.19)$$

В останній формулі вважаємо, що $Z_i(x)$ репрезентує власне семантичні параметри одиниці x у значенні $I_i(x)$, $U_i(x)$ – параметри синтактики, $V_i(x)$ – параметри прагматики; $\sigma_i(x)$, $\tau_i(x)$ та $\pi_i(x)$ є числовими коефіцієнтами, які репрезентують внесок (вагу, амплітуду) семантичних, синтактичних та прагматичних аспектів відповідно, у лексичне значення $I_i(x)$, причому:

$$\sigma_i(x) + \tau_i(x) + \pi_i(x) = 1; \sigma_i(x) \geq 0; \tau_i(x) \geq 0; \pi_i(x) \geq 0. \quad (1.20)$$

Таким чином, повний семантичний стан лексеми x набуває такого вигляду:

$$s(x) = \sum_i \beta_i \left[\sum_r \sum_l \alpha_r^i t(x^r) g_l^r(x^l) I(g; l; x^l) \right] \\ [\sigma_i(x)Z_i(x) + \tau_i(x)U_i(x) + \pi_i(x)V_i(x)], \quad (1.21)$$

в якій окремі параметри визначено формулами (1.12) – (1.20). Фіксуючи значення цих параметрів, одержуємо різні диференціальні семантичні стани лексеми x .

Слід зауважити, що операції множення та додавання елементів структури семантичних станів у формулах (1.12) – (1.21) наразі мають символічний характер. Адже поки що немає засобів для зіставлення цим елементам конкретних алгебраїчних структур, так що на даному етапі зазначені формули дають радше інтуїтивне уявлення про конструкцію семантичних станів, ніж цілком «строге» їх визначення. Але, незважаючи на це, аналіз формули (1.21) дозволяє зіставити семантичний стан лексеми відповідній словниковій статті тлумачного словника, надає достатньо однозначні інтерпретації складу і структури цієї словникової статті, а також і словника в цілому. І навпаки, з будови формули (1.21) можна вивести конструкції словникових статей для семантичних словників різних типів. Багатопараметричність цієї формули і змістовність її складових демонструє її гнучкість і великі можливості щодо експлікації та деталізації широкого кола лексико-граматичних і лексико-семантичних ефектів через конструктиви семантичних станів. Отже, викладений формалізм має значні потенції до узагальнень. Власне, вже зараз можна перейти до викладу будови абстрактної лексикографічної системи, спроможної моделювати й описувати семантику об'єктів, що мають набагато загальнішу інформаційну природу, ніж ті, якими оперує природна мова.

Розділ 2

ТЕОРІЯ ЛЕКСИКОГРАФІЧНИХ СИСТЕМ

Цей розділ присвячено викладові формалізованих принципів лексикографічного опису систем. Зазначені принципи ґрунтуються на інформаційній теорії систем [274], завдяки чому зазначений опис, побудований на засадах інформаційної інтерпретації, набуває рис набагато загальніших, ніж ті, що стосуються безпосередньо мовної системи і, фактично, стосується всіх об'єктів, де відбуваються інформаційні процеси з певними характеристиками та системними визначеннями. За типологією підхід до опису лексикографічних систем слід охарактеризувати як феноменологічний, оскільки він базується на найзагальніших інформаційних властивостях систем і не є наслідком їх конкретної будови. Водночас ми намагалися при розгляді інформаційних аспектів реальності бути максимально наближеними до такого кола систем, які є близькими до мовних або можуть бути кваліфіковані як такі. Очевидно, що для здійснення цієї настанови потрібно хоча б найзагальніше визначення мови – її системних рис, визначальних властивостей, які вирізняють мову серед інших інформаційних сутностей. На цьому шляху одним з авторів монографії (В. А. Широковим) було сформульоване фундаментальне для опису інформаційних об'єктів розглядуваного типу поняття лексикографічного ефекту в інформаційних системах, з викладу якого ми починаємо цей розділ.

2.1. Лексикографічний ефект в інформаційних системах

Процеси лексикографування як різновиду інтелектуальної діяльності та феноменологія словників, що виступають результатами цієї діяльності, не є постійними у часі величинами – вони еволюціонують відповідно до внутрішнього розвитку лінгвістичної науки та потреб практики. У певні

Історичні періоди саме зовнішні по відношенню до власних завдань мовознавства чинники – як це вже неодноразово було в історії науки – перетворюються на головну рушійну силу, яка визначає розвиток не тільки самої лексикографії як окремої ділянки лінгвістики, але й мовознавчої науки в цілому. Автори переконані, що відзначений розвиток позначиться і на прогресі інформаційної науки – упевненість у цьому надає аналіз напрямку руху інформаційного суспільства.

Справді, з появою комп'ютерних технологій виник безпрецедентний в історії світової цивілізації феномен – спілкування людини з неживою істотою за допомогою і через посередництво природної людської мови. Як би скептично не ставитися до розумових потенцій обчислювальної машини, сьогодні вже не відмовитися від факту, що мовні реакції сучасного комп'ютера в деяких випадках неможливо відрізнити від реакцій людини, а зважаючи на прогрес лінгвістичної технології протягом останніх двадцяти років, можна впевнено прогнозувати, що через деякий час комп'ютерами буде перебрано істотну частину мовної та лінгвістичної компетенції людини, що створить реальні передумови для побудови інформаційно-комп'ютерних систем на засадах природної мови.

Зазначена компетенція є важливою складовою частиною мовнорозумового апарату людини, а в ньому, як уже встановлено психо- та нейролінгвістикою, мовні структури нерозривно пов'язані зі структурами мислення. Яку роль тут відіграє словник? Як зауважив Р. Шенк [273], у людському мовнорозумовому апараті функціонують різні *словники* (які навіть більше нагадують енциклопедії), а дослідження протягом останніх десяти років, що ґрунтуються на принципах психолінгвістики, не лише підтверджують цю тезу, а й навіть проливають світло на можливу будову суб'єктивного лексикону людини.

Та обставина, що існує принципова можливість моделювання людської мови на неживих об'єктах, якими є комп'ютери, нагтовхує на думку, що в природі матерії також діють механізми, які мають з людською мовою спільні риси – отже, феномени, які можуть бути кваліфіковані як вияви мови, не обов'язково мусять бути виключною прерогативою людських (і взагалі живих) істот. У науці клас явищ та об'єктів, що характеризуються як мовні, відзначається великою різноманітністю. По-перше, це сама природна мова, яка існує тільки у формі національних мов (їх зараз у світі налічується близько шести тисяч і взаємовідношення між ними набувають подекуди драматичного характеру [152]). Існують також інші природні семіотично-семантичні системи, які з певною метафоричністю визначаються як мови – наприклад, мова генетичного коду. Крім того, створено

цілу низку артефактів мовної орієнтації: штучні мови типу есперанто, що імітують певні природні; формальні алгебро-алгоритмічні конструкції, які одержали назву формальних мов, останнім часом усе помітніше їхнє зближення з природними (в математичній лінгвістиці навіть існує термін майже природна мова). Про поширеність терміна мова в інформаційній науці свідчить терміносистема інформатики, яка містить такі поняття: мова програмування, інформаційно-пошукова мова, мова класифікаторів, мова запитів, мова опису даних, мова маніпулювання даними, мова опису предметної галузі, ціла низка так званих мов розмічування (SGML – Standard General MarcUp Language, HTML – HyperText MarcUp Language, XML – eXtended MarcUp Language, VRML – Virtual Reality MarcUp Language) і т. д. У 1960–1970 роки сформувалася ціла галузь інформатики – лінгвістичне забезпечення інформаційних систем.

Виняткову роль у культурі й цивілізації відіграють штучні інтерпретації природних мов – такими, зокрема, виступають їхні писемні варіанти (інколи, забуваючи про штучне походження письма, писемний варіант, який є, по суті, не більше ніж моделлю мови, називають просто – мовою). Усі штучні мовні конструкції адаптують ті чи інші ознаки і властивості природних мов, а їх зближення з останніми містить ознаки тих національних мов, носіями яких є автори відповідних штучних конструкцій.

У всіх названих мовних систем є спільні риси, для осмислення яких і побудови працездатної концептуальної схеми необхідне таке узагальнене уявлення про мову, яке можна було б застосувати до будь-якої з перелічених мовних систем, адже в сучасному інформаційному середовищі мова перестає бути виключною прерогативою людини, принаймні на технологічному рівні.

Що говорить про це лінгвістична наука? Вона твердить (висловлюючись, як і кожна фундаментальна наука, про свій предмет), що питання про визначення мови не просте, їх існує багато й вони різні – з огляду на ті чи інші аспекти цього багатогранного феномену. Незважаючи на їхню розмаїтість, узагальнюючи істотні риси, можна помітити, що більшість із них виявляється варіаціями на тему, запропоновану У. Чейфом [266]: мова – це система, яка в досить складний спосіб здійснює зв'язок між світом звуків і світом значень.

Значимо, що існує й багато інших дефініцій мови. Чимало з них ще з часів В. Гумбольдта ґрунтується на розмежуванні понять мови і мовлення. Після опублікування книги Ф. де Соссюра «Курс загальної лінгвістики» [253] це питання набуло широкого розголосу в лінгвістичних колах, причому діапазон думок коливався (і продовжує коліватися) у надзвичай-

но широких межах. Їх можна звести до трьох основних тверджень: 1) мовлення і мова протиставляються одне одному як цілком автономні об'єкти, що різняться між собою за сукупністю суттєвих ознак, унаслідок чого їх дослідження ведеться двома самостійними галузями науки – лінгвістикою мови та лінгвістикою мовлення; 2) мова і мовлення становлять єдиний об'єкт лінгвістики, а тлумачення відмінностей між ними покладається на методологію та зміст цієї комплексної науки; 3) між мовою і мовленням узагалі немає жодної різниці.

Л. В. Щерба, як відомо, виділяв три основних аспекти мовних явищ: під першим він розумів мовленнєву діяльність, під другим – словники і граматики, які виводяться на підставі актів мовлення і розуміння, що існували в певний історичний час серед людей певних груп, тобто мовні системи, під третім – сукупність того, що говорили або розуміли такі групи – мовний матеріал [299]. Він підкреслював зумовленість мовної діяльності складним мовним апаратом людини або психофізіологічною мовленнєвою організацією індивіда, яка: а) не може дорівнювати сумі мовленнєвого досвіду й повинна бути своєрідним його опрацюванням; б) може бути тільки психофізіологічною; в) є соціальним продуктом; г) слугує індивідуальним виявом мовної системи, що виводиться з мовного матеріалу; про характер цієї організації можна судити тільки на підставі мовленнєвої діяльності індивіда. Л. В. Щерба розмежував поняття механізму (мовленнєвої організації людини) і процесу (мовленнєвої діяльності); процесу і його продукту. Мовленнєву організацію він розглядає як єдність процесу й продукту; останній виступає індивідуальною системою концептів та стратегій використання індивідами у процесах мовлення й розуміння, котра й позначається як мова.

Не дискутуватимемо про відносну правильність тих чи інших поглядів, оскільки всі вони містять ті чи інші характерні ознаки феномену, для якого використовуємо спільну назву мова, й варіюються в різних комбінаціях у великому масиві лінгвістичних праць загального характеру. Опускаємо тут виклад, оскільки вважаємо, що в їхній основі – явно або неявно – також лежить відповідність між світом звуків і світом значень. Не беремо до уваги також визначень типу «мова + душа народу» та аналогічних, оскільки їх не вдається дослідити науковими методами.

Зазначимо, що для побудови продуктивної схеми (навіть мінімально формалізованої) використати визначення, про які йшлося, важко. Якщо світ звуків ми ще можемо якось локалізувати, то зі світом значень справа набагато складніша. Справді, де цей світ зосереджений? Як його дістати? Як ним оперувати? Взагалі, який зміст має твердження про існування

світу значень? Як зі світом значень співвідносяться світ змістів, світ зображень і багато інших світів? Отже, поєднання в одному визначенні означає, що дуже різняться за ступенем абстракції (світу звуків зі світом значень), викликає відчуття логічного розриву в цій дефініції та ставить більше запитань, ніж дає відповідей.

Проте і викладену, й інші спроби означення мови вважаємо корисними, оскільки кожна з них дає матеріал для узагальнення та абстрагування істотних рис такого універсального явища, яким є мова. Проаналізувавши їх і подивившись (прислухавшись) до будь-якого явища, яке ми інтуїтивно ідентифікуємо як мову з огляду на те, що вона робить і як вона працює, можемо зробити такий висновок: мова становить певний інструмент – своєрідну машину, яка забезпечує перетворення форми на зміст і навпаки. Але не довільної форми, а форми лінійних послідовностей певних дискретних об'єктів (знаків і знакових комплексів, звуків і звукових комплексів тощо).

Це твердження, не відрізняючись новизною¹, дає привід для заглиблення у феноменологію відношення «форма–зміст» (надалі паралельно вживатимемо для цього відношення аббревіатуру ВФЗ) і з'ясування таких його деталей, які б допомогли нам у абстрагуванні суттєвих рис мови. При цьому ми вважаємо, що в онтологічному вимірі ВФЗ не є апріорним іманентом об'єкта як «речі в собі», а такою його властивістю, що розкривається перед суб'єктом (дається) у процесі його взаємодії з об'єктом. Ми прагнемо до побудови формалізованого опису ВФЗ, адаптованого до створення спеціальної моделі даних, у якій були б представлені ефективні механізми рецепції істотних властивостей мови з технологічною орієнтацією на підтримку процесу створення словників та інших лінгвістичних продуктів. Для проведення аналізу деталей розгортання ВФЗ розглянемо діаграму, де символічно зображено процес сприйняття якогось об'єкта певним суб'єктом:

$$S : D \longrightarrow V(D). \quad (2.1)$$

Тут літерою D позначено щось із реального (або уявного) світу, що виступає в ролі об'єкта процесу сприйняття (спостереження, вивчення, уваги, переживання...) з боку певного S , яке ми вважаємо суб'єктом даного процесу; через $V(D)$ позначимо результат цього процесу. Зауважимо, що в ролі S може виступати людина, сконструйований людиною прилад, людино-машинна система або будь-що інше, наділене властивостями сприй-

¹ Різні аспекти відношення «форма – зміст» активно експлуатуються в мовознавстві з часів В. Гумбольдта, а потім і Ф. де Соссюра; у попередньому розділі ми також виклали усталені лінгвістичні погляди на це відношення.

няття й відчуття (відображення); S може бути і колективним суб'єктом – групою людей, соціальною спільнотою, етносом, нацією, народом, сукупністю народів або навіть людством у цілому.

Отже, не деталізуватимемо принципи і способи улаштування суб'єкта S , за винятком одного: S має дві визначальні властивості: а) сприйняття й відчуття зовнішніх сигналів (перцептивно-сенсорний апарат) і б) здатність до їх внутрішнього опрацювання (усвідомлення та інтерпретації). Таким чином, унаслідок фізичної, психічної, інтелектуальної та іншої обмеженості суб'єкта S уся сукупність властивостей об'єкта D для його сприйняття розподіляється на дві не дуже чіткі, неоднозначні, мінливі й не чітко відокремлювані частини. До першої з них залучаємо ті властивості D , які безпосередньо сприймаються перцептивно-сенсорним апаратом S , – позначимо цю частину через $F(D)$ і трактуватимемо її як сукупність властивостей D , співвіднесених із його формою з погляду суб'єкта S , який сприймає D . До другої частини включаємо властивості D , які безпосередньо не сприймаються перцептивно-сенсорним апаратом S , а відбиваються в ньому опосередковано через засоби внутрішнього опрацювання. Позначимо цю частину через $C(D)$ і розглядатимемо її як сукупність змістових властивостей D – знову-таки з погляду сприйняття того самого суб'єкта S . У зв'язку з цим діаграма (2.1) набуває такого вигляду:

$$D \xrightarrow{S_F} F(D) \xrightarrow{H} C(D), \quad (2.2)$$

де символом S_F позначено дію перцептивно-сенсорного апарату суб'єкта S , результатом якої є сукупність формальних (з погляду S) властивостей D ; символом H позначено механізм, який здійснює зв'язок між формою та змістом і забезпечує цілісність сприйняття об'єкта D суб'єктом S (якщо йому справді вдається забезпечити зазначену цілісність). Водночас, припустивши існування механізму, який дозволяє перехід від D до $C(D)$, – позначимо цей механізм через S_C – одержуємо таку трансформацію діаграми (2.2):

$$\begin{array}{ccc}
 D & \xrightarrow{S_F} & F(D) \\
 \downarrow S_C & & \swarrow H \\
 & & C(D)
 \end{array} \quad (2.3)$$

де, як бачимо, відбулася декомпозиція суб'єкта S на S_F та S_C , що реконструюють формальні та змістові властивості D відповідно.

Отже, в наведеній декомпозиції суб'єкт виступає таким, що має дуалістичну природу: він має механізм для безпосереднього (перцептивно-сенсорного – з огляду на форму речей) сприйняття об'єктивної дійсності D , а також наділений апаратом для інтелектуального опрацювання одержаної інформації шляхом витлумачення її змісту. Крім того, він наділений і певним інтерфейсом між цими двома способами опрацювання інформації, який у схемі представлено елементом H .

Не будемо абсолютизувати описану схему. Між $F(D)$ і $C(D)$ немає чіткої межі, як її взагалі немає між формою та змістом. Також майже ніяк не деталізувалися властивості S , хоч із загальних міркувань було здійснено його декомпозицію на S_F та S_C . Отже, такий підхід справді має ознаки феноменологічного, оскільки не спирається на припущення щодо можливої конструкції S і механізмів його функціонування. З цих міркувань можна стверджувати, що викладена схема є досить загальною – в неї не закладено жодних ангажментів. Єдиним, специфічним для мови, вважатимемо припущення, що $F(D)$ повинен мати лінійний характер, тобто зображатися лінійними послідовностями дискретних об'єктів, джерелом яких виступає певна скінченна множина.

З огляду на сказане навіть сама можливість існування такого явища, як мова, виступає наслідком фундаментальної властивості S бути суб'єктом, тобто тим, для котрого будь-що має свою зовнішню і внутрішню сторону (форму і зміст).

Відношення між цими різними аспектами сприйняття, символічно зображені величинами S_F , S_C , H , відзначаються великим розмаїттям, джерелом якого є фундаментально притаманні (тобто такі, яких у принципі не можна позбавитися) властивості сприймаючого суб'єкта S : мінливість, нерегулярність, різноманітність, обмеженість, нечіткість тощо.

Важливий аспект ВФЗ пов'язаний із властивістю концентрації уваги суб'єкта S на окремих фрагментах і форми і змісту. Це досягається настроюванням його перцептивно-сенсорної системи на деталі того, що він сприймає, внаслідок чого початкове ВФЗ модифікується: його певні змістові елементи набувають властивостей форми, а до формальних додаються нові, раніше не помічені деталі. До цього класу властивостей, зокрема, належать поняття внутрішньої і зовнішньої форми мовних одиниць.

Відзначимо ще одну рису процесу розгортання ВФЗ. Сучасна культурологія схильна трактувати її як символ постмодернізму, а саме – вплив суб'єкта на об'єкт, себто можливість зміни стану об'єкта D у процесі його

сприйняття (спостереження, дослідження...) суб'єктом S . Адже для того, щоб відбувся процес, символічно зображений діаграмами (2.1)–(2.3), у багатьох випадках необхідно активізувати об'єкт D , щоб він проявив свої властивості, якими цікавиться S . У класичній науковій парадигмі вважалося, що таке збудження об'єкта можна зробити скільки завгодно малим і знехтувати ним, вважаючи, що воно істотно не змінює стану об'єкта. Проте розвиток науки виявив, що насправді це, взагалі кажучи, не так.

Викладене вище виявляється співзвучним із зовсім іншими явищами та їх формальними моделями, які були побудовані незалежно, з іншого приводу і для інших цілей. Ідеться про визначення інформації А. М. Колмогоровим [173:213–223].

Уведення колмогорівської інформаційної міри, як відомо, має на меті уточнення поняття інформації, по-перше, без залучення ймовірнісного підходу, а по-друге, надання можливості застосування цієї міри до індивідуальних об'єктів.

Основна ідея підходу полягає в тому, що інформація про певний об'єкт вважається одержаною тоді, коли можливо відтворити (реконструювати) цей об'єкт (модель об'єкта) за деяким його скінченим описом (набором ознак). Побудова міри Колмогорова базується на таких фундаментальних поняттях, як алгоритм, машина Тьюринга, рекурсивна функція, і походить від ідей теорії складності обчислень (складності алгоритмів), яка, фактично, і є виток інтерпретації інформації як міри складності й структурованості систем, а також упевненості в універсальності такої характеристики як складність, оскільки будь-яка система, незалежно від її природи, характеризується певною складністю та має певну структуру, хоча б і тривіальну.

Відповідна, не переобтяжена деталями математична конструкція, яку ми тут подаємо, слідує за згаданою працею А. М. Колмогорова, формулюється у такий спосіб.

Розглянемо певну зліченну множину $X = \{x\}$. Вважатимемо, що існує взаємнооднозначна відповідність між X та множиною D двійкових слів, які починаються з одиниці – іншими словами, нехай задано бієктивне відображення:

$$n: X \rightarrow D, \quad (2.4)$$

таке, що кожному $x \in X$ однозначно відповідає певний $d = n(x)$, $d \in D$, і навпаки. Вважатимемо, що:

1) $n(x)$ – загальнорекурсивна функція на D . Позначимо через $l(d)$ довжину двійкового слова $d \in D$, тобто число нулів та одиниць, яке у ньому міститься. Тоді $l(n(x)) = l(x) + C$, де C – певна константа;

2) існує однозначне відображення $\chi: X^2 = X \times X \rightarrow X$, таке, що для $\forall x \in X, y \in X \exists z \in X$, що $z = \chi(x, y) \equiv (x, y)$ і $n(z) = n(x, y)$ є загальнонорекурсивною функцією від $n(x)$ та $n(y)$, причому

$$l(x, y) \leq C_x + l(y),$$

де константа C залежить лише від x .

Вважатимемо ізоморфізм (2.4) встановленим так, що множину X також розглядатимемо як множину двійкових слів.

Припустимо, що існує частково рекурсивна функція $\varphi(p, x)$, яка ставить у відповідність двійковому слову x двійкове слово y , причому $p, p \in D$ інтерпретується як алгоритм (або програма), яка переводить x на y :

$$p: x \rightarrow y, \quad (2.5)$$

а φ репрезентує при цьому метод (мову) програмування. Без втрати загальності вважатимемо, що p для даного x задається певним двійковим словом.

Позначимо:

$$K_\varphi(y|x) = \begin{cases} \min_p l(p), & \text{якщо } \varphi(p, x) = y \\ \infty, & \text{якщо не існує скінченного } p, \text{ такого, що } \varphi(p, x) = y. \end{cases} \quad (2.6)$$

Таким чином, $K_\varphi(y|x)$ є довжиною мінімальної програми p , яка переводить x в y при заданому методі програмування φ . Ця величина називається складністю y відносно x при даному φ . Звичайно, залежність величини складності від φ є невеликим описаного методу, але існує теорема [173:220], яка стверджує існування найкращого методу програмування A , такого, що для будь-якої частково рекурсивної функції φ справедлива нерівність:

$$K_A(y|x) \leq K_\varphi(y|x) + C_\varphi, \quad (2.7)$$

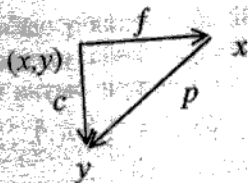
де стала C_φ залежить лише від φ і не залежить від x та y .

Величину $K_A(y) \equiv K_A(y|1)$, віднормовану відносно одиничного елемента $x = 1$, природно вважати складністю елемента y . При цьому кількість інформації в об'єкті x відносно об'єкта y визначається як різниця:

$$I_A(x|y) = K_A(y) + K_A(y|x). \quad (2.8)$$

Остання формула і визначає міру інформації – так звану алгоритмічну міру інформації Колмогорова.

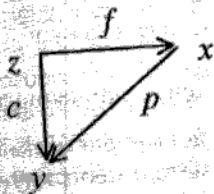
Формула (2.8) і весь описаний підхід визначення інформації через алгоритмічну складність можуть бути проінтерпретовані у дещо інший спосіб. З цією метою збудуємо трикутну діаграму:



(2.9)

де x та $y \in X$; елемент $(x, y) \in X^2$, причому завдяки існуванню відображення $\chi: X^2 \rightarrow X$, $\chi(x, y) = z$, $z \in X$. У діаграмі (2.9) відображення f та c здійснюють проєкції елемента (x, y) на перший та другий співмножники відповідно, причому справедливими залишаються формули (2.6)–(2.8) та інтерпретація складності й інформаційної міри.

Припустимо тепер, що об'єкт z , який досі у нас представляв образ декартового добутку (x, y) при відображенні χ , насправді є самостійним, до певної міри незалежним від x та y об'єктом зовнішнього світу. Це припущення дозволяє збудувати таку діаграму:



(2.10)

де $z \in Z$, Z – множина-джерело об'єктів z , $x \in X$, $y \in Y$. Тоді відображення f та c визначають певні інтерпретації об'єкта z , і, крім того, відображення p інтерпретує x через y . Природно припустити, що об'єкт x відбиває формальні властивості об'єкта z , а y – його змістові властивості, причому зв'язок між формою та змістом бере на себе p . Вимога мінімальності p тут є цілком зрозумілою, адже витлумачення форми (а його результат – це власне і є зміст!) не повинно містити зайвих, випадкових щодо x (а також і z) елементів. Побудована у такий спосіб конструкція, яка ґрунтується на трійці об'єктів (z, x, y) та трійці відображень (f, c, p) , що формують діаграму (2.10), допускає природну інтерпретацію як комплекс, який є носієм ВФЗ.

Бачимо, що ВФЗ у такий своєрідний спосіб виявляється зашифрованим в алгоритмічній теорії складності та інформації, а значить, це відношення є тісно пов'язаним з інформаційними процесами; навіть більше – воно, на наше переконання, є їхнім системотвірним відношенням. Відночас воно є й основним системотвірним відношенням природної мови, що наполегливо і переконливо підтверджують відомі лінгвістичні теорії.

Отже, мовнорозумові об'єкти, процеси, конструкції та структури мають інформаційну природу, до якісного і кількісного аналізу якої цілком правомірне застосування формалізму Колмогорова. Очевидна схожість, спорідненість діаграми (2.3) з визначенням інформації за Колмогоровим (див. формули (2.5–2.6), а також подальші міркування, які привели до побудови діаграми (2.10)), наводять на думку про те, що в основі і одного, і другого лежать ті самі закономірності. Сама форма представлення інформаційної міри свідчить про певний процес, результатом якого стала генерация алфавіту – знакової системи представлення об'єкта. Розгортання відображень (f, c, p) у діаграмах (2.9)–(2.10), яким зіставляються елементи ВФЗ з діаграми (2.3), спонукає до порівняння останніх зі складовими інформаційних процесів, які в теорії Колмогорова зводяться до математичних відношень (алгоритмів, рекурсивних функцій, ...), визначених на дискретних множинах.

Описана ситуація здається настільки загальною, що дозволяє зробити висновок: в основі будь-якого процесу, явища, об'єкта, системи тощо лежить певна дискретна сукупність – назовемо її сукупністю елементарних інформаційних одиниць (надалі паралельно вживатимемо абревіатуру ЕЮ), у визначенні якої ключову роль відіграє поняття лексикографічного ефекту в інформаційних системах, суть якого така.

Загальною ознакою всіх процесів інформаційного обміну є перетворення інформації з однієї форми в іншу, причому сучасні природничі теорії однозначно підтверджують принципово дискретний (квантовий) характер процесу взаємодії та обміну і, отже, принципову дискретизацію процесу опису дійсності. Зазначена дискретизація має принаймні одну спільну для всіх відомих процесів ознаку, на наш погляд, фундаментального характеру: спостерігаючи та узагальнюючи поведінку різних систем, доходимо висновку, що в процесі еволюції (динаміки, саморозвитку тощо) системи будь-якої природи в її структурі індукується певна підсистема відносно сталих дискретних сутностей (підсистема порядку), які відіграють роль її ЕЮ, тому всі інші феномени системи являють собою певним способом організовані комбінації ЕЮ.

Зазначена підсистема має властивості, споріднені з властивостями лексичної системи природної мови: вона генерує в своїй структурі щось на зразок тезаурусу і граматики з притаманними цим конструкціям знаковістю, значенням, змістом, поліморфністю, омонімією (ізоморфізмом), синонімією (гоморфізмом), полісемією (багатозначністю), еліпсисом, метонімією, комбінованістю і т. д.; вона є носієм як плану вираження, так і плану змісту. Розгортання, взаємодія, взаємовпливи та взаємозумовленість

обох зазначених планів у системі елементарних інформаційних одиниць підлягають певним закономірностям, причому багатомірову контраверзу натуралістів і конвенціоналістів ми схильні розв'язувати позитивно: існують приклади систем обох типів — як таких, де план вираження детермінований планом змісту, так і таких, зв'язок між якими має конвенціональний характер.

Сукупностям елементарних інформаційних одиниць властива субстанційність, як і будь-яким агрегаціям, спричиненим об'єктивними процесами (у даному випадку це лексикографічний ефект), унаслідок чого зазначені сукупності, як правило, мають відносну стабільність своїх характеристик, що забезпечує їхню локалізацію у відповідних областях системних параметрів.

Описана сукупність явищ становить зміст лексикографічного ефекту. Можна стверджувати, що при вивченні будь-яких предметних галузей фахівці фактично займаються дослідженням лексикографічних ефектів, які відбуваються в них або є для них характерними. Таким чином, лексикографічний ефект повинен розглядатися не лише з феноменального боку, а й з методологічного, оскільки має певний потенціал операціональності, стимулюючи в процесі моделювання тих чи інших систем встановлення й визначення відповідних комплексів ЕІО, враховуючи, репрезентуючи та конкретизуючи їхні властивості. У цій іпостасі концепція лексикографічного ефекту виступає як метод абстрагування даних.

Оскільки лексикографічний ефект виражається у представленні певного континуального універсуму через посередництво дискретних множин, виникає приваблива можливість його формалізації через так званий парадокс Сколема, відомий у математиці як теорема Левенгейма-Сколема [223]. Остання полягає у тому, що при деяких обмеженнях існує певний ізоморфізм між незліченною та зліченною множинами; у певному смислі, потенційна нескінченність може інтерпретуватися скінченним способом, по суті, будується фінітна модель нескінченності. Теорема Левенгейма-Сколема говорить, що будь-яка несуперечлива теорія першого порядку, яка має незліченну модель, має також і зліченну модель. Це означає, що якщо деяка структурована множина задана зліченною множиною правил, то існує зліченна множина (тобто власна підмножина множини натуральних чисел), на якій можна побудувати точну модель цієї структурованої множини, де будуть виконані усі вихідні аксіоми. Отже, має місце те ж саме представлення нескінченного об'єкта через скінченний об'єкт, що містить у собі всю інформацію про нескінченний. У такий спосіб теорема Ле-

венгейма-Сколема справді відіграє роль формального корелята лексикографічного ефекту, принаймні певних його різновидів [138].

У подальшому комплекс ЕЮ системи D , що розвинувся внаслідок лексикографічного ефекту Q , позначатимемо через $I^Q(D)$ або $I(D)$, коли посилення на тип лексикографічного ефекту несуттєве. Система елементарних інформаційних одиниць, будучи носієм цілої низки властивостей, має певну структуру, що є їхнім виявом. Зокрема, в будь-якій системі ЕЮ завжди виділяється ядро – певна власна підсистема $I_0^Q(D) \subseteq I^Q(D)$ – і визнається породжувальна процедура π :

$$\pi : I_0^Q(D) \rightarrow I^Q(D). \quad (2.11)$$

Трійку $(I^Q(D), I_0^Q(D), \pi)$ також ототожнюватимемо з системою ЕЮ і використовуватимемо це позначення поряд з $I^Q(D), I(D), I_0^Q(D), I_0(D)$ як еквіваленти, вважаючи, що породжувальна процедура π визначена, відома і зрозуміла з контексту.

Підсумовуючи зміст викладеної феноменології, констатуємо, що процеси, аналогічні описаним в цьому підрозділі, відбуваються у всіх достатньо складно побудованих соціотехнічних системах і взагалі у системах будь-якої природи, в яких діють джерела, перетворювачі та споживачі інформації, а отже, відбуваються аналоги перцептивно-сенсорних актів та мовнорозумових процесів. Це дає додаткове обґрунтування справедливості висновку про існування та універсальність лексикографічного ефекту в інформаційних системах з характеристиками знаковості, значення, змісту й поліморфності, що споріднюють його з природною мовою, якій також притаманні відзначені властивості [251]. Вказані чинники підсилюють бажання побудови якомога формальніше визначених структур, які є носіями лексикографічних ефектів, та їхнього застосування до опису мовних феноменів.

2.2. Структура і архітектура лексикографічних систем

Базовим конструктивом побудов такого типу ми визначаємо так звані лексикографічні системи (надалі паралельно вживатимемо скорочення: Л-системи). Поняття Л-системи, на наше переконання, є фундаментальним, а його означення ґрунтується на викладеній феноменології лексикографічного ефекту. Лексикографічні системи уособлюють досить загальний тип формалізованих конструкцій у ряду таких, як моделі даних, формальні граматики, формальні системи, канонічні числення у скінченних алфавітах тощо.

Зазначимо, що з частинними випадками (або реалізаціями) лексикографічних систем у науці й техніці оперують дуже давно: ними виступає велика кількість різноманітних інформаційних систем, баз даних та знань, до кола яких потрапляють і всі традиційні словники та комп'ютерні словникові системи.

Якщо говорити про машинні словники, то вони спроможні ефективно виконувати свої функції тільки тоді, коли їхня структура достатньо повно відображає форму та зміст одиниць мови, що є об'єктами лексикографування. Тенденція до відтворення цієї повноти насправді спостерігається лише у тих випадках, коли проектування інформаційно-лінгвістичних систем ґрунтується на глибокому вивченні феноменології мови, яка сама підказує вибір адекватного апарату, а також конструкції відповідних моделей. Незважаючи на те, що метою інформаційної науки є витлумачення предметної галузі (у нашому випадку – лінгвістичних фактів) мовою моделей даних, самі типи та конструкції цих моделей повинні впливати з предметної галузі і якомога точніше враховувати специфіку мовних явищ. Виходячи зі сформульованих положень, було збудовано структурну теорію лексикографічних систем, засновану на феноменології лексикографічного ефекту, послідовне застосування якої надало можливість для необхідних системних узагальнень і створення методології побудови лексикографічних моделей.

Історично початковим пунктом аналізу, результатом якого стало формулювання теорії лексикографічних систем, виявилось дослідження значного числа структур реально існуючих традиційних словників, їхнє узагальнення й побудова відповідних моделей. З метою деталізації та інформаційно-лексикографічної конкретизації було проведено дослідження загальних структуротвірних ефектів та елементів лексикографічних систем, які абстрагуються з традиційних словників, перетворюючись на елементи інфологічних моделей лексикографічних систем загального положення. Цей шлях привів до встановлення поняття структури лексикографічних систем.

Очевидно, що структура традиційних словників не випадкова, оскільки в ній сконцентровано багатовіковий досвід цілих поколінь лексикографів. Тому вона, як правило, є вільною від суб'єктивних смаків та уподобань розробників інформаційних систем. Досвід лексикографування як різновиду інтелектуальної діяльності (в міру його нагромадження) від систематизації власне лінгвістичних фактів (і навіть ще вужче – від систематизації даних про лексичні одиниці) поступово поширюється й на сис-

тематизацію даних про світ, знання про який, у свою чергу, зосереджені у природній мові як цілісній інформаційній системі.

Універсальність явища лексикографічного ефекту спричиняє неодноразово відзначену нами тенденцію до лексикографування будь-якого мовного феномену – саме цей факт пояснює побутування в лексикографічній практиці прикладів створення словників, у яких лексикографуються навіть і такі одиниці мови, які не мають безпосереднього вербального вираження. Так, спроба лексикографування синтаксичних структур зроблена, наприклад, у праці Г. А. Золотової [162], у вступі до якої зазначається: «Як фізичний світ навколо нас складається з елементарних частинок, найдрібніших з відомих частинок матерії, так і синтаксична будова нашої мови організується різноманітними, але регулярними комбінаціями елементарних, або мінімальних, одиниць, далі неподільних на синтаксичному рівні. У лінгвістиці на сучасному етапі її розвитку визріла потреба в осмисленні поняття елементарних синтаксичних одиниць, з яких, як стає все очевидніше, будуються й усі інші, складніші конструкції». І далі: «Синтаксемою названо мінімальну, далі неподільну семантико-синтаксичну одиницю російської мови, яка виступає одночасно як носій елементарного смислу і як конструктивний компонент складніших синтаксичних побудов, яка, отже, характеризується певним набором синтаксичних функцій» [162:3–4]. Відзначимо очевидну аналогію (подекуди майже текстуальний збіг) з нашим формулюванням лексикографічного ефекту, простір дії якого, очевидно, є незрівнянно ширшим.

Аналогічні спроби лексикографування семантичних структур [75] не лише відбивають загальну тенденцію до лексикографічного опису всіх мовних явищ, а й відповідають потребам практики щодо розроблення найдосконаліших систем лінгвістичного забезпечення.

Із викладеного випливає методологічна коректність включення одиниць будь-якого мовного рівня до складу елементарних інформаційно-лексикографічних одиниць тієї чи іншої лексикографічної системи. У такий спосіб лексикографуються семантичні, синтаксичні, когнітивні та інші структури, що, як правило, не мають безпосереднього вербального представлення в системі природної мови. До такого типу робіт близькі й праці зі створення словників: ідеографічних, дієслівного керування, еквівалентів слів, фразеологізмів тощо. До останніх двох типів словників примикає ціла низка можливих лексикографічних праць, поки ще не створених, але які теоретично мають цілковите право на існування [147]. У згаданій праці наведено пропозиції до створення понад 50 різних словників, у яких одиницями лексикографування (елементарними інформаційними одиницями

щодо відповідних, подекуди вельми екзотичних лексикографічних ефектів) виступають, наприклад, звернення, етикетні фрази, гоноративи (висловлювання ввічливості), гумілятиви (вирази хамства), стимули і реакції (підтакування, згоди, заперечення, спростування) тощо.

Вивчення різноманітних структур існуючих традиційних словників дозволяє зробити певні узагальнення, які можна не лише покласти в основу теоретичної лексикографічної схеми, а й використати при проектуванні конкретних інформаційних систем лінгвістичного напрямку й створенні відповідного програмного забезпечення. Оскільки в лексикографії вже давно відбулося розмежування понять *словник* та *список слів, перелік, індекс, інвентар*, словник як абстрактна лексикографічна система обов'язково наділений структурою, що містить мінімум дві необхідні частини: реєстрову (ліву) та інтерпретаційну (праву), що є виявом ВФЗ. Саме наявність другої — носія змістового компонента ВФЗ — відрізняє словник від звичайного списку слів. Але він має глибшу структуру, яка відбивається в будові реєстрової та інтерпретаційної частини словника в цілому і його окремих словникових статей, а також у структурі міжстатейних та міжсловникових відображень. Відтак словник — це спеціальний вид тексту, в якому в систематизованому та структурованому вигляді подається опис лексики певної мови (або сукупності мов). Однак його природно розглядати і як специфічний об'єкт техніки, а саме — інформаційну систему, де через поліграфічне виконання означено певні лінгвістичні ефекти за допомогою шрифтових виділень, позиційного розміщення, спеціальних позначок тощо, які відіграють роль ідентифікаторів відповідних інформаційних змінних — елементів метамови словника. Складність будови словника полягає ще й у тому, що не всі елементи його структури явно означені вказаним вище способом. У структурі реальних словників, як правило, існує велика кількість неявних структурованих елементів, виявлення яких часто є досить складним завданням. Процес абстрагування словникової (лексикографічної) структури становить своєрідне розшифрування, реконструкцію того лексикографічного ефекту, який спричинив її утворення і розгортається з використанням кількох положень, сформульованих спершу в мовознавстві, але по суті таких, що мають загальносистемний характер.

Побудова структурної моделі лексикографічних (словникових) систем орієнтується на багатоаспектне представлення знакової природи лексичних одиниць як найкомпактніших і найінформативніших у природній мові. З позицій теорії лексикографічного ефекту це означає виділення в досліджуваній мовній системі підсистеми її елементарних інформаційних одиниць і визначення множини їх системно-структурних параметрів.

Наступний момент полягає в урахуванні дихотомічної структури кожної елементарної інформаційної одиниці (й повної їх сукупності), що відбивається в багатовимірному співвідношенні форми та змісту, носієм якого є визначений клас елементарних інформаційних одиниць.

Багатоаспектність представлення знакової природи одиниць мови в традиційних словниках (або елементарних інформаційних одиниць в загальних лексикографічних системах) забезпечується врахуванням лінгвістичних і когнітивних особливостей об'єктів лексикографування – залежно від типу словника й глибини характеристики лексикографічного ефекту, який виступає предметом дослідження в кожному конкретному випадку. В інформаційно-лексикографічній моделі із зазначеними особливостями зіставляється певна множина комплексів даних та / або знань.

Зауважимо, що в мовному (мовленнєвому) потоці онтологічна природа мови виступає неподільною на окремі складові частини, які є в концептуальних представленнях. З цього випливає прагнення до створення інтегральних словників, а отже, необхідність використання комплексних (інтегрованих) моделей мовних явищ, тому при розробленні комп'ютерних систем опрацювання мови постає завдання створення таких формалізованих моделей, які були б налаштовані на ефективне представлення інтеграційних процесів і водночас урахували б специфіку лінгвістичних об'єктів. Таким чином, критерій багатоаспектності репрезентації знакової природи одиниць мови дозволяє побудувати комплексні, інтегровані моделі даних, придатні до поєднання концептуальних представлень різних за своєю природою мовних явищ.

Дихотомічність структури елементарних інформаційних одиниць в інформаційно-лексикографічній моделі (подібно до того, як це робиться в більшості традиційних словників) виявляється в структурній організації лексикографічної системи й впливає з фундаментальних положень сучасної лінгвістики, яка оперує поняттями форми і змісту, внутрішньої та зовнішньої форм мовних одиниць, феноменологія яких глибоко простежена на мовному матеріалі.

Як відзначав В. М. Русанівський [236], мова має дуалістичну функцію: з одного боку, це матеріальна основа, на яку мислення спирається в процесі свого функціонування, а з другого – матеріал, у якому воно фіксується, стаючи dokonаним фактом. Об'єктами дослідження складників мовнорозумового потоку є як фізична (матеріальна), так і змістова (ідеальна) сторони. Отже, звукову субстанцію мовлення можна вважати його формою, а інформаційні властивості – змістом. З огляду на цю обставину, звукова реалізація мовлення може поділятися на елементи різного ступеня

агрегованості – інтонаційно цілісні одиниці (інтонеми), комбінації голосних і приголосних (склади), власне голосні та приголосні (звуки) і т. д. Цей процес необмежений, оскільки виділення та класифікація складників звуків мовлення залежать від цілого ряду чинників, у тому числі й від прогресу акустики, фонології тощо. Фізичний процес мовлення належить до незворотних і (як і багато інших акустичних явищ) дисипативних процесів. Зазначені властивості фізичної субстанції мовлення разом з властивостями мовленнєвого апарату визначають її зовнішні інформаційні характеристики.

У свою чергу, писемна форма мови моделює її усну форму, тому взагалі справедлива послідовність: <модель дійсності – мислення> → <модель мислення – усна мова> → <модель усної мови – писемна мова>. Оскільки наведені моделі фізично реалізуються в єдиній системі (пов'язаній з індивідуумами, соціальними спільнотами, системами культури тощо), цілком природні та необхідні їх взаємодія та взаємовпливи. Таким чином, писемний варіант мови так само виступає у функції моделі мислення і моделі дійсності.

Безпосереднє буття мови у вигляді мовленнєвої діяльності, а також писемності та інших способів фіксації мовних актів на фізичних носіях, відмінних від природномовних, репрезентує властивість мови мати зовнішню форму. Остання взагалі можлива завдяки здатності мови бути представником феноменальної сторони дійсності, і оскільки мова сама становить певну дійсність, у ній закладені потенції для позначення (представлення) самої себе.

Система, яка виступає репрезентантом феноменальної сторони дійсності, повинна бути й певним чином організованою. Оскільки відмінність між явищем і сутністю є відносною, а між феноменальною і сутнісною дійсністю немає чіткої межі, її не повинно бути й у мові як моделі дійсності. Цей факт реалізується у властивості слова мати внутрішню форму, яка пов'язана з представленням ноуменальної частини його власного буття саме в системі мови. Зовнішня і внутрішня форми, таким чином, взаємопов'язані й разом становлять форму слова на противагу його змісту – сумі конкретних значень.

Усе це дає підстави для твердження, що ВФЗ (включаючи й уявлення про внутрішню та зовнішню форму мовних одиниць) мають загальний характер, репрезентують універсальну властивість елементарних інформаційних одиниць, що індукуються в процесі розвинення будь-якого лексикографічного ефекту, і будучи формалізованими у вигляді відповідних моделей даних, спроможні утворити субстрат моделей інформаційних систем довільної природи та походження, а для мовно-орієнтованих моде-

лей даних вони взагалі є обов'язковими. Зазначені поняття, на наш погляд, мають потенціал конструктивності, оскільки будь-який зміст існує лише у певній формальній оболонці, що дозволяє застосовувати уніфікований підхід до побудови їх репрезентантів у науковій теорії.

Розглянемо певну систему D , концептуальний опис якої представимо у вигляді певної лексикографічної системи. Оскільки нас цікавлять насамперед лінгвістичні факти, системою D у нас виступатиме саме природна мова, сукупність природних мов або певна підсистема (певний аспект) природної мови.

Відповідно до викладеного системі D властива складна ієрархія лексикографічних ефектів. Так, для природномовної системи можна навести приклади цілої низки лексикографічних ефектів, результатом яких стало виділення з мовленнєвого потоку окремих фонем, складів, морфем, слів (словоформ), лексем, словосполучень, речень тощо. Усі названі структурні одиниці виступають як складники відповідних класів елементарних інформаційних одиниць стосовно тих чи інших типів природномовних лексикографічних ефектів.

У подальшому як лексикографічну систему розглядатимемо спеціальне інформаційне середовище, в якому розвивається (реалізується) певний лексикографічний ефект (або певна сукупність лексикографічних ефектів).

Для побудови практично корисної схеми моделювання наведених явищ необхідно визначити набір інформаційних конструктивів, які специфікують структурні елементи Л-систем, дозволяючи розробляти конкретні застосування. У свою чергу, це спричиняє необхідність побудови конструктивної теорії Л-систем – в її основу покладено лексикографічну модель даних, розроблену в працях [276; 288], результатами і позначеннями яких ми скористаємося у подальшому викладі.

Відповідно до інформаційної інтерпретації процесів сприйняття [288] визначимо результат рецепції суб'єктом S класу елементарних інформаційних одиниць $I^Q(D)$ у вигляді певної множини $V(I^Q(D))$ – множини описів одиниць, що належать до класу $I^Q(D)$; ця множина є результатом процесу:

$$S : I^Q(D) \rightarrow V(I^Q(D)), \quad (2.12)$$

тому для кожного елемента $x \in I^Q(D)$ однозначно визначено його опис $V(x)$ як елемент множини $V(I^Q(D)) : V(x) \in V(I^Q(D))$; $Sx = V(x)$. Отже, логічно припустити, що $V(I^Q(D))$ має вигляд об'єднання:

$$V(I^Q(D)) = \bigcup_{x \in I^Q(D)} V(x). \quad (2.13)$$

Згідно з інформаційною концепцією представлення опису системи ЕЮ, кожний $V(x)$ зображається у вигляді слова (тексту) в певному скінченному алфавіті $A = \{a_1, a_2, \dots, a_n\}$, тобто скінченної послідовності символів з A (надалі слова в алфавіті A називатимемо A -словами). Наприклад, якщо ми розглядаємо Словник української мови, то його алфавіт A складається з таких елементів:

- звичайний алфавіт української мови (великі й малі літери);
- знаки пунктуації;
- арабські цифри;
- римські цифри;
- символи пробілу та абзацу;
- спецсимволи ($//$, Δ , \blacktriangle , \diamond , \blacklozenge , ...);
- типи шрифтів тощо.

У такий спосіб опис $V(x)$ будь-якої елементарної інформаційної одиниці x , $x \in I^Q(D)$, представляється A -словом такого вигляду:

$$V(x) = v_1(x)v_2(x)\dots v_{k(x)}(x), v_i(x) \in A, i = 1, 2, \dots, k(x), k(x) \geq 1, \quad (2.14)$$

де кожна A -літера $v_i(x)$ береться з алфавіту A . Відзначимо, що довжина $k(x)$ A -слова $V(x)$ залежить від x – тобто від того, яким обрано елемент $x \in I^Q(D)$. Формула (2.14), за визначенням, подає достатньо повний, у певному смислі вичерпний, опис елементарної інформаційної одиниці x в даній лексикографічній системі. Між класом ЕЮ $I^Q(D)$ і множиною описів $V(I^Q(D))$ за допомогою відображення S встановлюється певний ізоморфізм. Іншими словами, множина описів $V(I^Q(D))$ є певною власною підмножиною множини $W(A) : V(I^Q(D)) \subset W(A)$, причому $W(A)$ — множина всіх слів скінченної довжини над A , тобто послідовностей вигляду $v_1v_2\dots v_q$, $q < \infty$, $v_i \in A$, $i = 1, 2, \dots, q$. Вважаємо, що слово нульової довжини – 0 також належить до $W(A) : \forall a \in W(A) \exists 0 \in W(A)$, причому $a * 0 = 0 * a = a$, де «*» – конкатенація. Замкненість відносно операції конкатенації, тобто вимога: $a, b \in W(A) \Rightarrow \exists c \in W(A), c = a * b$, а також асоціативність відносно неї: $\forall a, b, c \in W(A) \Rightarrow a * b * c = (a * b) * c = a * (b * c)$ перетворює $W(A)$ на напівгрупу з напівгруповою операцією «*» і одиничним елементом 0 .

Вибір алфавіту A , через який зображено $W(A)$ та $V(I^Q(D))$, відповідно алгебраїчній традиції тут не обґрунтовуватимемо і не конкретизуватимемо, проте зауважимо, що його генерація є наслідком певного лексикографічного ефекту, який розвивається в системі мовлення (акустичного) та його інформаційно-графічної інтерпретації. Якщо ми розглядаємо звичайні словники, то природною є інтерпретація A -слова $V(x)$ як тексту словникової статті з реєстровою одиницею x .

Взагалі структура напівгрупи є досить бідною, а конструкція $W(A)$ – занадто широкою для ефективного виявлення в ній характерних властивостей мовних систем. Для досягнення цієї мети необхідне запровадження додаткових припущень та обмежень, за допомогою яких у структурі $W(A)$ виділяються підструктури, характерні саме для природної мови. Це досягається в такий спосіб.

Оскільки кожний $V(x)$, як зазначалося, є адекватним й однозначним репрезентантом (описом) відповідного елемента x із системи $I^Q(D)$, в його структурі повинні бути з достатньою повнотою відображені властивості цього елемента. Враховуючи лінійний характер об'єкта $V(x)$, котрий зображається лінійною послідовністю символів з A , доходимо висновку: єдиним природним джерелом його структури може виступати лише певна множина A -підслів та певні відношення між її елементами. A -підслова в описі $V(x)$ визначаються як A -слова, складені з тих символів алфавіту A , що містяться в розглядуваному описі $V(x)$ і розташовуються в A -підслові у порядку, індукованому порядком розташування літер у самому описі. Очевидно, що множина всіх A -підслів A -слова довжини n (тобто A -слова, складеного з n A -літер) містить 2^n елементів. Позначимо множину всіх A -підслів A -слова $V(x)$ через $B[V(x)]$.

Структура на множині описів вводиться у такий спосіб. Припустимо, для всіх описів $V(x)$ існує єдине правило, за яким з будь-якого A -слова $V(x)$ можна виділити множину A -підслів $\beta(x) = \{\beta_i(x)\}$ з такими властивостями:

- елемент x належить до множини $\beta(x)$;
- весь опис $V(x)$ є елементом множини $\beta(x)$;
- правило, за яким виділяються елементи множини $\beta(x)$, є єдиним для всіх $V(x)$.

Описаним способом з будь-якого $V(x)$ виділяється множина $\beta[V(x)]$ величин (A -підслів) $\beta_i(x)$ такого вигляду:

$$\beta[V(x)] \equiv \{\beta_i(x), i = 1, 2, \dots, q\} \subseteq B[V(x)], \quad (2.15)$$

де $B[V(x)] = \{v_{i_1} v_{i_2} \dots v_{i_p}, 1 \leq i_1 < i_2 < \dots < i_p \leq k(x), p = 1, 2, \dots, k(x)\}$, причому:

$$\begin{aligned} v_{ij} &\in \{v_{1(x)}, v_{2(x)}, \dots, v_{k(x)}(x)\}; x \in \beta[V(x)]; \\ V(x) &\in \beta[V(x)], \beta_{k(x)} \neq \beta_{m(x)} \text{ при } k \neq m. \end{aligned} \quad (2.16)$$

Покладемо за визначенням:

$$\beta[V(I^Q(D))] = \bigcup_{x \in I^Q(D)} \beta[V(x)]. \quad (2.17)$$

Очевидно, що $V(I^Q(D)) \in \beta[V(I^Q(D))]$. Позначимо

$$\beta_i = \bigcup_{x \in I^Q(D)} \beta_i(x), \quad i = 1, 2, \dots, q, \quad \text{а також } \beta = \bigcup_i \beta_i. \quad (2.18)$$

Зрозуміло, що $\beta \equiv \beta[V(I^Q(D))]$. Відзначимо, що деякі з елементів $\beta_i(x)$, $i = 1, 2, \dots, q$, можуть бути порожніми для певних $x \in I^Q(D)$; у цьому випадку вони випускаються у формулах (2.15)–(2.18).

Через $\sigma[\beta]$ позначимо певну структуру, визначену на β , а отже, на $V(I^Q(D))$, – надалі називатимемо $\sigma[\beta]$ макроструктурою $V(I^Q(D))$; обмеження $\sigma[\beta]$ на $V(x)$: $\sigma[\beta]S_{V(x)} \equiv \sigma(x)$ породжує мікроструктуру $V(x)$. Активне формулювання цього факту полягає у встановленні процедури (оператора, процесу...) σ , який породжує на β структуру $\sigma[\beta]$:

$$\sigma : \beta \rightarrow \sigma[\beta]. \quad (2.19)$$

На β можлива генерація цілої низки неізоморфних структур $\sigma[\beta]$, у ролі яких можуть виступати будь-які з відомих моделей даних (ієрархічна, мережева, реляційна, об'єктно-реляційна та ін.), логіко-математичні моделі (зокрема, логічні числення типу логіки предикатів), конструкції формальних граматик тощо.

Одним із можливих механізмів формування структури може бути такий. Збудуємо таблицю:

β_1	β_2	...	β_q
$\beta_1(x_1)$	$\beta_2(x_1)$...	$\beta_q(x_1)$
$\beta_1(x_2)$	$\beta_2(x_2)$...	$\beta_q(x_2)$
...
$\beta_1(x_M)$	$\beta_2(x_M)$...	$\beta_q(x_M)$

Деякі з елементів $\beta_i(x_j)$, очевидно, можуть бути порожніми, тому довжини стовпчиків таблиці, взагалі кажучи, є різними. Величини β_i , $i = 1, 2, \dots, q$ інтерпретуватимемо як атрибути (імена атрибутів), а набори $Dom \beta_i \equiv \{\beta_i(x_1), \beta_i(x_2), \dots, \beta_i(x_M)\}$, $i = 1, 2, \dots, q$ як домени цих атрибутів. Тоді структура $\sigma[\beta]$ може бути реалізована у вигляді певної реляційної алгебри $R[\beta]$, визначеної на декартовому добутку:

$$\prod_{i=1}^q Dom \beta_i = \beta^{\otimes}. \quad (2.20)$$

Іншими словами, якщо структура σ ототожнюється з певною реляційною алгеброю R над β^{\otimes} , то трійка $\{V(I^Q(S)), \beta, R[\beta]\}$ представляє собою не що інше, як реляційну модель, а п'ятірка $\{I^Q(S), D, V(I^Q(S)), \beta, R[\beta]\}$ задає деяку об'єктно-реляційну модель [176]. При цьому $I^Q(S)$ представляє

клас об'єктів моделі; β_i інтерпретуються як атрибути (імена атрибутів) з доменами $Dom \beta_i$, елементами котрих служать $\beta_i(x)$, $x \in I^Q(S)$. Зрозуміло, що окремими доменами є сама множина $\{x\}$ (для скорочення викладу без подальшої деталізації ототожнюватимемо елемент x як такий, що належить до класу $I^Q(S)$, з його ім'ям у $V(x)$), і $V(I^Q(S))$ – як множину $\{V(x)\}$. Реляційні відношення відповідних арностей визначаються як завжди – у вигляді певних підмножин множини β^{\otimes} . Їх кортежами виступають елементи вигляду:

$$(\beta_{i_1}(x_{j_{i_1}}), \beta_{i_2}(x_{j_{i_2}}), \dots, \beta_{i_r}(x_{j_{i_r}})), \quad (2.21)$$

$$i_1 < i_2 < \dots < i_r; x_{j_{i_m}} \in I^Q(S), m = 1, 2, \dots, r.$$

Реляційне числення в цій моделі визначається звичайним чином (див., наприклад, роботу Дж. Ульмана [262]).

Завдяки можливості інтерпретації $I^Q(D)$ як класів, елементами яких виступають об'єкти будь-якого походження, природною є об'єктно-орієнтована інтерпретація моделі. Відношення між елементами класу $I^Q(D)$ індукуються системою унарних відношень $r[\beta_1]$ на $\beta_1 = \{x\}$ та відображенням:

$$\Delta : V(I^Q(D)) \rightarrow I^Q(D); \Delta r[\beta_1]. \quad (2.22)$$

У такий спосіб комплекс $I^Q(D)$ постає представником онтологічної природи дійсності, яка підлягає моделюванню, у той час як $V(I^Q(D))$, β , $\sigma[\beta]$ репрезентують її концептуальну сторону.

Подальший розгляд концентруватиметься навколо ВФЗ, носієм яких є комплекс $I^Q(D)$ і які розвиваються й реалізуються в середовищі $\{I^Q(D), S, V(I^Q(D)), \beta, \sigma[\beta]\}$. При цьому сукупність властивостей і якостей комплексу $I^Q(D)$, згідно з викладеним вище, розподіляється на дві не цілком чіткі й не дуже виразно відокремлювані частини. Зауважимо, що в концептуальній схемі, реалізованій в описі $V(I^Q(D))$, вони повинні бути відокремлені; інакше кажучи, необхідною умовою коректності її побудови є існування процедури, яка здійснює таке відокремлення. Усе це можна відобразити в комутативній діаграмі:

$$\begin{array}{ccc}
 & V(I^Q(D)) & \\
 F \swarrow & & \searrow C \\
 \Lambda(I^Q(D)) & \xrightarrow{H} & P(I^Q(D))
 \end{array} \quad (2.23)$$

де $FV(I^Q(D)) = \Lambda(I^Q(D))$; $CV(I^Q(D)) = P(I^Q(D))$; $\Lambda(I^Q(D)) \cap P(I^Q(D)) = \emptyset$, причому $H \circ F = C$, символом « \circ » позначено композицію відображень;

$$\Lambda(I^Q(D)) = \bigcup_{x \in I^Q(D)} \Lambda(x); P(I^Q(D)) = \bigcup_{x \in I^Q(D)} P(x). \quad (2.24)$$

На $\Lambda(I^Q(D))$ і $P(I^Q(D))$ індукуються макроструктури:

$$F\sigma[\beta] = \lambda[\beta] \text{ і } C\sigma[\beta] = \rho[\beta] \quad (2.25)$$

та відповідні мікроструктури:

$$\lambda[\beta]S_{V(x)} \equiv \lambda(x); \rho[\beta]S_{V(x)} \equiv \rho(x). \quad (2.26)$$

як обмеження $\lambda[\beta]$ та $\rho[\beta]$ на $V(x)$.

Відзначимо, що діаграма (2.23), тобто об'єкти $V(I^Q(D))$, $\Lambda(I^Q(D))$, $P(I^Q(D))$ і відображення F , C , H , є загальним, нефакультативним елементом лексикографічної моделі і фактично будується незалежно від структури $\sigma[\beta]$. Походження та зміст її складових елементів зовсім інші. А саме: $\Lambda(I^Q(D))$ співвідноситься з тією частиною опису $V(I^Q(D))$, яка – в певному розумінні – представляє форму $I^Q(D)$, а $P(I^Q(D))$, відповідно, зіставляється з тією частиною опису $V(I^Q(D))$, що відповідає за зміст $I^Q(D)$. Наведена специфікація практично підтверджує думку про те, що ВФЗ є універсальним, притаманним об'єктам будь-якого походження, оскільки вони залучаються до процесу взаємодії з суб'єктом, який їх сприймає або досліджує.

Визначення 1. Вісімка об'єктів $\{I^Q(D), S, V(I^Q(D)), \beta, \sigma[\beta], F, C, H\}$ визначає елементарну лексикографічну модель даних, а її конкретна реалізація – елементарну лексикографічну систему. Інколи для скорочення, коли не виникатиме різночитань, будемо позначати через $V(I^Q(D))$ цілу елементарну лексикографічну систему.

Зауважимо, що будь-який елемент β_i (або будь-яка їх сукупність), який належить до структур β , $\sigma[\beta]$, $\lambda[\beta]$, $\rho[\beta]$, може бути інтерпретований як елементарна Л-система. Звідси одержуємо можливість виокремити в структурі початкової елементарної Л-системи низку інформаційно-лінгвістичних підструктур, які тлумачимо як окремі Л-системи. Перевизначивши у такий спосіб структуру вихідної Л-системи, одержуємо Л-модель та Л-систему загального положення (не елементарну); вона представлена у вигляді об'єднання певної кількості елементарних Л-систем з можливими відображеннями та зв'язками між ними. Таким чином, Л-система загального положення має вигляд графа $G = \{V = \{V_i\}; R = \{R_{kl}\}\}$, де $V = \{V_i\}$ – множина вершин, якими виступають елементарні Л-системи V_i , що входять до G , а $R = \{R_{kl}\}$ – множина ребер графа G , R_{kl} , що поєднує V_k та V_l .

Зокрема, ніщо не заважає розглядати $\Lambda(I^Q(D))$ і $P(I^Q(D))$ як окремі, автономні елементарні Л-системи, а це уможливило таку побудову:

$$\begin{array}{ccc}
 V(I^Q(D)) = (\Lambda(I^Q(D)) \equiv \Lambda_0(I^Q(D))) & \xrightarrow{H_0} & P_0(I^Q(D)) \equiv P(I^Q(D)) \\
 \swarrow F_{01}^\Lambda & & \swarrow F_{01}^P \\
 \Lambda_{01}^\Lambda(I^{Q_1}(D)) & \xrightarrow{H_{01}^\Lambda} & P_{01}^\Lambda(I^{Q_1}(D)) \\
 \searrow C_{01}^\Lambda & & \searrow C_{01}^P \\
 & & \Lambda_{01}^P(I^{Q_2}(D)) \xrightarrow{H_{01}^P} P_{01}^P(I^{Q_2}(D))
 \end{array} \quad (2.27)$$

Звернімо увагу на зміну типу лексикографічного ефекту на другому поверсі – замість Q тепер маємо Q_1 й Q_2 відповідно. Отже, приходимо до комплексів об'єктів $I^{Q_1}(D)$ та $I^{Q_2}(D)$. Продовжуючи цей процес, одержуємо рекурсивне розвинення лексикографічної системи $V(I^Q(D))$:

$$\begin{array}{ccc}
 & V = (\Lambda_0; P_0) & \\
 & \swarrow & \searrow \\
 \Lambda_0 & & P_0 \\
 \swarrow & & \swarrow \\
 \Lambda_{01}^\Lambda & & \Lambda_{01}^P \\
 \swarrow & & \swarrow \\
 & & \dots
 \end{array} \quad (2.28)$$

Назвемо цей процес *рекурсивною редукцією лексикографічної системи*. Він нагадує своєрідний інформаційний «мікроскоп», що виявляє все тонші деталі структури лексикографічної системи й індукує структуру, схожу на фрактальну.

Надалі позначатимемо процес рекурсивної редукції Л-системи $V(I^Q(D))$ через $RR\downarrow[V(I^Q(D))]$. У визначення цього процесу входять, характеризації всіх операторів F, C, H на всіх наявних рівнях рекурсивної редукції разом із результатами їх дії, а також усі макро- і мікроструктури σ, λ, ρ .

Викладена конструкція становить зміст загального визначення лексикографічної моделі даних:

$$\{I^Q(D), S, V(I^Q(D)), \beta, \sigma [\beta], RR\downarrow[V(I^Q(D))]\} \quad (2.29)$$

та лексикографічної системи:

$$\{I^Q(D), S, V(I^Q(D)), \beta, \sigma [\beta], RR\downarrow[V(I^Q(D))], \Sigma\}, \quad (2.30)$$

де символом Σ позначено її архітектуру як інформаційної системи.

Визначення окремих елементів моделі у формулах (2.29)–(2.30) представлено формулами (2.23)–(2.28); архітектура Σ зазвичай обирається трирівневою, такою, що відповідає ANSI/X3/SPARK (або просто ANSI/SPARK) [306]. У дещо іншому формулюванні аналогічну архітектуру було запропоновано у звіті [309], де також запроваджено три рівні абстракції даних: концептуальний – фізичний – представлень. З певними уточненнями можна вважати справедливою відповідність названих трьох рівнів концептуальному, внутрішньому та зовнішньому рівням архітектури ANSI/SPARK. Слід згадати також працю Цикритзиса і Клуга [317], яка є неформальним вступом до переглянутої версії звіту [306]. Три рівні абстракції представлені у великій кількості існуючих баз даних. Основні складники архітектури ANSI/SPARK використовуватимемо в такій інтерпретації:

$$ARCH_LS = \{CM, EXM, INM; \Phi, \Psi, \Xi\}, \quad (2.31)$$

де символом CM позначено концептуальну модель лексикографічної системи LS . Символом $EXM = \{exM\}$ – множину її зовнішніх моделей, які відповідають CM , а $INM = \{inM\}$ – відповідна множина її внутрішніх моделей. Через $\Phi = \{\phi\}$ позначено множину відображень CM в EXM :

$$\phi : CM \rightarrow exM, \text{ де } exM \in EXM; \quad (2.32)$$

відповідно $\Psi = \{\psi\}$ – множина відображень CM в INM :

$$\psi : CM \rightarrow inM, \text{ де } inM \in INM; \quad (2.33)$$

$\Xi = \{\xi\}$ – множина відображень INM в EXM :

$$\xi(inM) = exM. \quad (2.34)$$

При цьому ми зупиняємося на такій інтерпретації елементів архітектури.

Концептуальна модель має такі властивості:

1. *Семіотичність.* Концептуальна модель реалізована в середовищі певної знакової системи. Її побудова ґрунтується на спільному розгляді предметної галузі, галузі мислення і знакової галузі.

2. *Семантичність.* В категоріях моделі відображаються об'єкти та зв'язки між ними, суттєві з точки зору адекватного опису знань про предметну галузь. Модель не повинна залежати від логічної, фізичної та зовнішньої репрезентації даних.

3. *Однозначність.* Концептуальна модель описує предметну галузь згідно з принципом однозначного іменування, згідно з яким кожен знак моделі має одне значення, один смисл. Омонімія і полісемія у визначенні

елементів моделі є знятою. Використання кваліфікацій і контекстних механізмів не допускається.

4. *Несуперечливість*. Для кожного стану концептуальної моделі межі між її категоріями є абсолютними: при класифікації чи описі об'єктів предметної галузі кожному з них відповідає однозначно визначена множина станів, які попарно не перетинаються.

5. *Інтегрованість*. У концептуальній моделі СУМ-20 суміщаються уявлення різних фахівців про предметну галузь. Суперечливість цих уявлень фіксується за допомогою засобів обмеження цілісності й усувається через посередництво спеціальних процедур.

6. *Типізованість*. Всі концептуально визначені властивості об'єктів моделі повинні мати інтерпретацію на певних типах даних.

7. *Алгоритмізованість*. Всі об'єкти концептуальної моделі, а також зв'язки, відображення та операції над ними повинні мати скінченний опис, тобто мати інтерпретацію у вигляді алгоритмів скінченної складності над конструктивними об'єктами.

Отже, концептуальна модель (концептуальний рівень представлення) предметної галузі – це знакова, семантична модель, у котрій в однозначно-му, скінченному та несуперечливому вигляді інтегруються уявлення різних фахівців про предметну галузь.

У внутрішній моделі (внутрішньому рівні представлення) визначаються типи, структури та формати представлення, зберігання та маніпулювання даними, алгоритмічна база та операційно-програмне середовище, в яке занурюється концептуальна модель при її реалізації.

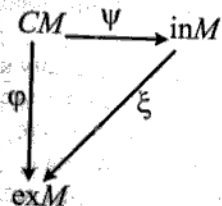
Зовнішня модель (зовнішній рівень представлення) відображає погляди кінцевих користувачів і, отже, прикладних програмістів на інформаційну систему. В ній реалізується система засобів, які дозволяють користувачеві здійснювати дозволені контакти й маніпулювання даними, представленими у внутрішньому рівні.

Одній концептуальній моделі може відповідати декілька внутрішніх та зовнішніх моделей, тому відображення $\Phi : CM \rightarrow EXM$ та $\Psi : CM \rightarrow INM$, взагалі кажучи, не ін'єктивні. Але відображення $\phi : CM \rightarrow exM$, та $\psi : CM \rightarrow inM$ будуються ін'єктивними (та зазвичай не бієктивними). Визначимо множину Ξ відображень INM в EXM :

для $\forall inM \in INM$ та $\forall exM \in EXM \exists \xi \in \Xi$ таке, що:

$$\xi(inM) = exM. \quad (2.35)$$

При цьому відображення ϕ, ψ, ξ будуються в такий спосіб, що діаграма:



(2.36)

є комутативною: $\xi \circ \psi = \phi$. Вимога комутативності цієї діаграми є суттєвою, оскільки гарантує узгодженість між усіма рівнями архітектури системи. У свою чергу:

$$CM = \{Ob(LS); RelOb(LS); Mor(Ob(LS), RelOb(LS));$$

$$Res(Ob(LS), RelOb(LS); Mor(Ob(LS), RelOb(LS)))\},$$

де $Ob(LS)$ – множина об'єктів та категорій LS ; $RelOb(LS)$ – множина зв'язків (відношень) між об'єктами та категоріями LS ; $Mor(Ob(LS), RelOb(LS))$ – множина можливих операцій (процесів) над $Ob(LS)$ та $RelOb(LS)$; $Res(Ob, RelOb, Mor(Ob, RelOb))(LS)$ – множина обмежень цілісності. Всі названі елементи мають однозначну інтерпретацію в термінах лексикографічної моделі даних $\{I^Q(S), d, V(I^Q(S)), \beta, \sigma[\beta], RR \downarrow [V(I^Q(S))]\}$.

Внутрішня модель лексикографічної системи містить такі елементи:

$$INM = \{D(t, s, f), ALG(D(t, s, f)), OS, PL\},$$

де $D(t, s, f)$ – множина даних, специфікована за типами, структурами та форматами t, s, f , відповідно, за допомогою яких у внутрішньому рівні представляються елементи CM ; $ALG(D(t, s, f))$ – множина алгоритмів (процесів) обробки та маніпулювання даними; OS – множина операційних платформ та PL – множина мов програмування (включаючи і процедурні мови типу СКБД – системи керування базами даних), на яких реалізовано $D(t, s, f)$ та $ALG(D(t, s, f))$.

Зовнішня модель лексикографічної системи представляється у вигляді:

$$EXM = \{IF, SC, FUNC, PROC, APR\},$$

де IF – інтерфейс системи; SC – множина сценаріїв; $FUNC$ – множина функцій; $PROC$ – множина допустимих процесів, APR – множина прикладних програм.

Конкретну інформаційно-лінгвістичну реалізацію лексикографічної системи у певному комп'ютерному середовищі надалі називатимемо лексикографічною базою даних (ЛБД).

2.3. Лексикографічні структури і словники

У застосуванні до традиційних словників та словникових комплексів лексикографічні системи набувають простої та прозорої інтерпретації. Справді, алфавіт Л-системи ототожнюється зі знаковою системою словника (включаючи й спецсимволи), клас ЕЮ $I^Q(D)$ – з множиною речетрових одиниць (об'єктів лексикографування – $\{x\}$), множина описів $V(I^Q(D)) = \{V(x)\}$ – з множиною словникових статей, де заголовні одиниці пробігають множину $\{x\}$, а $\Lambda(x)$ та $P(x)$ – з лівими та правими частинами відповідних словникових статей і т. д.

Застосування конструктивів Л-систем надає засоби для далекосяжних узагальнень традиційних завдань лексикографії, а багато з її нерозв'язаних проблем набувають свого природного вирішення у парадигмі Л-систем.

Наведемо для прикладу проблему упорядкування словникових статей за різними критеріями, яка у традиційному словнику не має природного розв'язку. У Л-системі зазначена проблема зводиться до задання системи класифікацій на множині $I^Q(D)$, яка служить основою для побудови відповідного пошукового апарату і реалізується відповідними засобами внутрішньої та зовнішньої моделі Л-системи. Найпростіша з таких класифікацій – абетка – породжує так зване лексикографічне упорядкування множини $I^Q(D)$ і, отже, всього словника. Морфемна класифікація, яка полягає у виділенні класів слів з однаковими основами, веде до гніздового принципу, який є характерним для словників словотвірного типу, а також використовується у двомовних перекладних словниках. Граматична класифікація може породити цілу гаму словників. Цей перелік можна продовжити. Зрозуміло, що традиційний паперовий словник природно упорядковується лише за одним із класифікаційних принципів (хоча існують і спроби поєднання різних принципів класифікації в одному словникові). Зовсім інші можливості відкриваються для загальних лексикографічних систем та їх реалізацій – комп'ютерних словників, пошуковий апарат яких може одночасно включати всі згадані, а також цілу низку інших класифікацій.

Визначимо у структурі $\sigma[\beta]$ Л-системи $V(I^Q(D))$ підмножину A спеціального типу, елементи A ($A \in A$) якої будемо називати *автоморфізмами* Л-системи $V(I^Q(D))$. Смісл цих елементів полягає в тому, що вони забезпечують внутрішні відображення $V(I^Q(D))$, тобто відображення:

$$A : V(I^Q(D)) \rightarrow V(I^Q(D)) \quad (2.37)$$

спеціального характеру, ще точніше – відображення між окремими словниковими статтями $A : V(x) \rightarrow V(y)$ для різних x та y . У конкретних лексикографічних працях автоморфізм A може, зокрема, констатувати наяв-

ність відсилкових словникових статей типу, наприклад, таких: x див. y . Вказаний автоморфізм визначає таке відображення словникових статей: $V(x) \rightarrow V(y)$. Його ідентифікатором є, як правило, деяке відсилкове псевдослово (у наведеному прикладі – див. y), яке зіставляє словниковій статті $V(x)$ її відповідник $V(y)$. Але будова автоморфізма A може бути складнішою, ніж у цьому прикладі.

По-перше, довжина низки відсилань може бути більшою за одиницю, тобто мати ланцюгово чи рекурсивно розгортальний характер:

$$V(x) \rightarrow \{V(x')\} \rightarrow \dots \rightarrow \{V(x'')\} \rightarrow \dots$$

Крім того, відображення $V(x) \rightarrow V(y)$ може репрезентувати цілий пучок відсилань. Це реалізується, наприклад, тоді, коли словникова стаття $V(x)$ має таку будову: x, x', x'', \dots див. y, y', y'', \dots

У цьому випадку в одній словниковій статті $V(X)$ визначено пучок відображень:

$$V(x) \rightarrow V(y); V(x') \rightarrow V(y'); V(x'') \rightarrow V(y'') \dots$$

Наприклад, словникова стаття в СУМ-11:

УГВИНТИТИ, УГВИНТИТИСЯ, УГВІНЧЕНИЙ, УГВІНЧУВАТИСЯ див. **вгвинтити, вгвинтитись** і т. д.

репрезентує пучок відсилань:

$$V(\text{УГВИНТИТИ}) \rightarrow V(\text{вгвинтити}),$$

$$V(\text{УГВИНТИТИСЯ}) \rightarrow V(\text{вгвинтитися})$$

і т. д.

Елементи множини автоморфізмів A не обов'язково повинні задаватися в явному вигляді. Більше того, встановлення словникових автоморфізмів, як правило, не формалізується, оскільки є досить складним завданням, пов'язаним з розкриттям внутрішніх закономірностей і прихованої структури (симетрії) Л-системи. Одну з таких прихованих симетрій, а саме, приховану симетрію дієслівної підсистеми Л-системи академічного тлумачного Словника української мови, буде досліджено у наступному розділі.

Множини відображень H та A породжують макроструктуру Л-системи. До макроструктурних елементів Л-системи відносимо також $F\sigma[\beta] = \lambda[\beta]$ і $S\sigma[\beta] = \rho[\beta]$, визначені формулою (2.25). Локальні відображення (вони будуються як обмеження відповідних макроструктур на $V(x)$): $HS_{V(x)}$, а також $\lambda[\beta]S_{V(x)} \in \lambda(x)$; $\rho[\beta]S_{V(x)} \in \rho(x)$ – два останні визначено формулою (2.26) – визначають мікроструктуру, яка відображає у неяв-

ному вигляді семантику предметної галузі, що є об'єктом даної конкретної Л-системи.

Встановлення і визначення мікроструктур лексикографічних систем дозволяє формалізувати, а у багатьох випадках й автоматизувати процес побудови структур відповідних словникових баз даних, що надає структурному підходові значних переваг при проектуванні елементів лінгвістичного забезпечення інформаційних систем. У третьому розділі докладно описано конкретний приклад застосування теорії та методології лексикографічних систем до побудови лексикографічної бази даних академічного тлумачного Словника української мови на основі його тексту.

Нижче ми вводимо деякі поняття, які дозволяють формально визначити структури, що індукуються на лексикографічних системах відображеннями спеціального вигляду. Ці відображення не тільки індукують природні структури на конкретних лексикографічних системах, але й дають певний інструментарій для породження структурної класифікації лексикографічних систем і методологію класифікації словників. З використанням цих відображень вдається сформулювати поняття близькості на множині слів (ззагалі, одиниць будь-якого рівня) – вживаємо термін *псевдотопологія* – і навіть відстані (*псевдовідстані*) між словами.

Розглянемо деяку елементарну Л-систему $V(I^W(L)) = \{V(x)\}$. Будь-яку її словникову статтю V_i можна представити у вигляді:

$$V_i = x_0^i \pi_1 \xi_1^i \pi_2 \xi_2^i \pi_3 \dots \pi_n \xi_n^i, \quad (2.38)$$

$$\bigcup_i \bigcup_j \xi_j^i = I^W(L), \quad (2.39)$$

де $I^W(L)$ – множина слів мови L , що містяться у всіх словникових статтях Л-системи $V(I^W(L))$; π_i – роздільники між словами (знаки пунктуації, службові позначки, спецсимволи, скорочення тощо; вони узагальнюються під назвою *mark-up symbols*). Отже, довільна словникова стаття V_i подається у вигляді об'єднання:

$$V_i = \partial V_i \cup M_i, \quad (2.40)$$

де вжито такі позначення: $\partial V_i \equiv x_0^i$ – граничний елемент словникової статті, її заголовне слово; M_i – внутрішня частина словникової статті:

$$M_i \equiv \pi_1 \xi_1^i \pi_2 \xi_2^i \pi_3 \dots \pi_n \xi_n^i. \quad (2.41)$$

Таким чином, будь-яка словникова стаття є об'єднанням внутрішньої частини та границі, а весь словник представляється у вигляді об'єднання множини внутрішніх частин з множиною границь. Позначимо:

$\partial V = \bigcup \partial V_i$ – границя елементарної Л-системи $V(I^W(L))$;

$M = \bigcup M_i$ – внутрішня частина елементарної Л-системи $V(I^W(L))$.

Отже:

$$V = \partial V \cup M. \quad (2.42)$$

Визначення 2. Лексикографічна система (2.42) називається замкненою, якщо для $\forall \xi \in I^W(L) \exists V(x_0^a) \equiv V_\alpha \in V$, що $x_0^a \equiv x_\xi$, де x_ξ – вихідна форма слова ξ .

Це визначення означає, що всі слова, наявні у словникових статтях Л-системи, є водночас заголовними словами певних словникових статей.

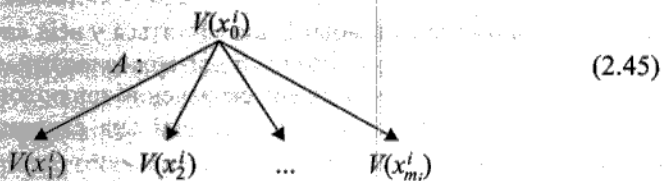
Визначимо автоморфізм A деякого спеціального типу в такий спосіб. Назвемо множину слів:

$$\langle V_i \rangle \equiv (x_1^i, x_2^i, \dots, x_{m_i}^i); m_i \leq n_i \quad (2.43)$$

кортежем словникової статті V_i , якщо вона являє собою множину канонічних (словникових) форм до слів $\xi_1^i, \xi_2^i, \dots, \xi_{m_i}^i$, відповідно. До кортежу необов'язково включаються всі слова з M_i . При розгляді конкретних прикладів, зокрема, при вивченні структури лексикографічних систем тлумачного типу, лексикографічна доцільність спонукає до обмеження кортежу лексемами, що входять тільки до формул тлумачення (словникових дефініцій), причому з випущенням з них так званих стоп-слів (таких, що не дають суттєвого внеску в семантику). Визначимо автоморфізм A за допомогою формули:

$$A: V(x_0^i) \equiv V_i \rightarrow \{V(x_k^i), k = 1, 2, \dots, m_i\}. \quad (2.44)$$

Графічна репрезентація формули (2.44) має такий вигляд:



Формули (2.44), (2.45) означають, що словниковій статті $V(x_0^i)$ із заголовним словом (границею) x_0^i ставляться у відповідність словникові статті $V(x_k^i), k = 1, 2, \dots, m_i$ із заголовними словами $x_k^i, k = 1, 2, \dots, m_i$, відповідно, якщо, зрозуміло, вони взагалі наявні у $V(I^W(L))$. Рекурсивно довизначимо дію оператора A на $V(x_1^i), V(x_2^i), \dots, V(x_{m_i}^i)$, і далі – на результатах його застосування до $V(x_1^i), V(x_2^i), \dots, V(x_{m_i}^i)$ і т. д., аж допоки об'єкти $V(x_j^i), i, j = 1, 2, \dots$ не почнуть повторюватися. Через $V^A[x_0^i]$ ($V^A[x_0^i] \subseteq V$) позначимо множину словникових статей $\{V(x_0^i), V(x_1^i), V(x_2^i), \dots, V(x_{m_i}^i), \dots\}$, одержану в результаті визначеної вище дії оператора A .

Визначення 3. Множина $V^A[x_0^i]$ називається $A[x_0^i]$ -підсловником словника $V(I^W(L))$, якщо $A V^A[x_0^i] = V^A[x_0^i]$.

Таким чином, A -підсловник $V^A[x_0^i]$ є інваріантною множиною у $V(I^W(L))$ відносно дії A .

Визначення 4. Елементи $V(x_0^i), V(x_1^i), V(x_2^i), \dots, V(x_{m_i}^i), \dots$ A -підсловника $V^A[x_0^i]$ називатимемо A -еквівалентними елементами.

Останнє визначення стає зрозумілим, якщо врахувати, що множина відображень A , яка породжує A -підсловник $V^A[x_0^i]$, індукує відношення еквівалентності на множині його словникових статей; його позначатимемо символом $EV^A[x_0^i]$. Позначимо через:

$$W = V \setminus EV^A[x_0^i] \quad (2.46)$$

фактормножину в V по відношенню $EV^A[x_0^i]$. Звідси: $V = W \cup V'$. Згідно з визначенням V' :

$$A V' = V' A^M W = V' \text{ для деякого } M \geq 0.$$

Тоді говоримо, що V являє собою напівпрямую суму словників W і V' :

$$V = W \triangleright V'. \quad (2.47)$$

Словник V з такою структурою будемо називати A -нерозкладним.

Визначення 5. Словник $V(I^W(L))$ називається A -незвідним (цілком незвідним), якщо в ньому немає власних A -підсловників.

З останнього визначення випливає, що коли V є A -незвідним словником, то для довільних $x, y \in I^W(L) \exists N \geq 0$, що $V(y) \subseteq A^N V(x)$.

Визначення 6. Словник V називається цілком A -звідним, якщо його можна подати у вигляді:

$$V = \bigcup_i V^i, \text{ причому } V^i \cap V^j = \emptyset \text{ при } i \neq j,$$

де V^i — A -незвідний словник. У цьому випадку будемо говорити, що V розкладається в пряму суму A -словників V^i :

$$V = \sum_i \oplus V^i. \quad (2.48)$$

Розглянемо деякий A -незвідний словник V :

$$V = \bigcup_{x_0^i \in S_0} V(x_0^i),$$

де $V(x_0^i)$ — словникова стаття з заголовним словом x_0^i . Автоморфізм A індукує відображення $S_0 \rightarrow S_0$, тобто він визначає деяке відображення множини заголовних слів в себе. Вказане відображення визначається у такий спосіб: будемо вважати, що коли:

$$A: V(x_0^i) \equiv V_i \rightarrow \{V(x_k^i), k = 1, 2, \dots, m_i\}, \text{ то}$$

$$A: x_0^i \rightarrow \{x_k^i, k = 1, 2, \dots, m_i\}.$$

З викладеного випливає

теорема 1. Для A -незвідного словника V не існує A -інваріантних підмножин в S_0 . Отже, для довільних $x, y \in S_0 \exists N(x, y) \geq 0$ таке, що $A^N x = y$.

Іншими словами, шлях $A^N x$ при достатньо великих N проходить через всі точки множини S_0 . Як наслідок справедлива

теорема 2. Повний A -шлях на графі $G^A(S_0)$ завжди замкнений.

Позначимо: $\inf N(x, y) = \rho(x, y)$.

Визначення 7. Число $\rho(x, y)$ називається A -псевдовідстанню від слова x до слова y .

Число $\rho(x, y)$ показує, за яку мінімальну кількість кроків можна дійти від слова x до слова y , застосовуючи алгоритм A .

Нехай множина слів:

$$(x_1^i, x_2^i, \dots, x_{m_i}^i; m_i \leq n_i) = \tau^i \equiv \tau(x_0^i)$$

є кортежем словникової статті V_i , тобто являє собою множину вихідних форм до слів $\xi_1^i, \xi_2^i, \dots, \xi_{m_i}^i$, відповідно. Очевидно, що x_0^i також належить до $\tau(x_0^i)$.

Визначення 8. Назвемо $\tau(x_0^i)$ замкненим околom точки x_0^i .

Визначення 9. Множина

$$\{\emptyset, \tau(x_0^i), i = 1, 2, \dots, \text{Card } S_0\} \quad (2.49)$$

називається (V, A) -псевдотопологією елементарної Л-системи.

Поняття псевдотопології спроможне відіграти важливу роль у теорії лексикографічних систем, оскільки за допомогою цього поняття виникає можливість формалізації поняття близькості елементарних інформаційних одиниць (слів). А саме, перегин околів $\tau(x_0^i) \cap \tau(x_0^j)$ визначає ступінь близькості лексем x_0^i та x_0^j , який випливає з лексикографічної системи.

За допомогою автоморфізмів A можна визначити ще деякі корисні структури на Л-системах. Робиться це у наступний спосіб.

Розглянемо ланцюг заголовних слів певних словникових статей тлумачного словника, який будується через застосування A -автоморфізмів:

$$x_0 \xrightarrow{A_1} x_1(x_0) \xrightarrow{A_2} x_2(x_1) \rightarrow \dots \rightarrow x_{n-1}(x_{n-2}) \xrightarrow{A_n} x_n(x_{n-1}). \quad (2.50)$$

Процес побудови ланцюга такий. Обирається словникова стаття із заголовним словом x_0 . Вона, очевидно, має будову $V_0 = \partial V_0 \cup M_0$, де елемент ∂V_0 є в точності x_0 , а внутрішня частина $M_0 \equiv \pi_1 \xi_1^0 \pi_2 \xi_2^0 \pi_3 \dots \pi_r \xi_r^0$ має

кортеж $(x_1^0, x_2^0, \dots, x_{r_0}^0)$. З цього кортежу обирається певний елемент $-x_1(x_0)$, який у результаті застосування процедури A_1 стає заголовним словом словникової статті V_1 . Вибір елемента $x_1(x_0)$ не випадковий. Доцільно, щоб він перебував у певному гіперонімічному відношенні до x_0 (рід-вид, частина-ціле, комплекс-елемент тощо). Наступні кроки побудови такого ланцюга, тобто знаходження слів:

$$x_2(x_1), \dots, x_{n-1}(x_{n-2}), x_n(x_{n-1}),$$

які є результатами застосування операторів A_2, A_3, \dots, A_n , відповідно, є аналогічними. Зауважимо, що ми не вимагаємо, щоб на кожному кроці побудови ланцюга його елементи обирались тільки в одному, фіксованому відношенні до попередника, тобто не вимагаємо, щоб відношення між словом та його попередником у ланцюгові було сталим для всього ланцюга. Назвемо побудовані в такий спосіб ланцюги слів вигляду (2.49) *гіперланцюгами* Л-системи V . Гіперланцюги, для яких відношення між сусідніми словами є сталим упродовж всього гіперланцюга, називатимемо *однорідними*.

Розглянемо процес побудови гіперланцюгів на прикладі словникової статті із СУМ-20 із заголовним словом $x_0 =$ абажур. Його лексичне значення задається дефініцією «Частина світильника, звичайно у вигляді ковпака, призначена для зосередження і відбиття світла та захисту очей від його впливу», вихідні форми слів якої належать до кортежу $\langle V_0 \rangle$. За $x_1(x_0)$ виберемо слово *світильник*, яке, очевидно, перебуває зі словом *абажур* у відношенні *частина-ціле*. Одним зі значень цього слова є «Прилад для освітлення – електричний або у вигляді лампи, плошки, куди наливають олию і вставляють тніт», з якого через дію автоморфізму A одержуємо: $A_2[x_1(x_0)] = x_2(x_1)$, де $x_2(x_1) =$ *прилад*. Слово *прилад* перебуває зі словом *світильник* у відношенні *рід-вид*. Продовжуючи цей процес, одержуємо такий гіперланцюг:

$$Z_1 = \text{абажур} \rightarrow \text{світильник} \rightarrow \text{прилад} \rightarrow \text{пристрій} \rightarrow \text{обладнання} \rightarrow \text{прилад}.$$

Як бачимо, на п'ятому кроці цього процесу гіперланцюг «зациклився», тобто знову повернувся до слова *прилад*, яке вперше з'явилось на третьому кроці. Такі гіперланцюги називатимемо *гіперциклами*. Для замкнених Л-систем, як випливає з теореми 2, будь-який достатньо довгий гіперланцюг є гіперциклом. Наведемо ще один приклад гіперциклу:

$$Z_2 = \text{абажур} \rightarrow \text{ковпак} \rightarrow \text{накривка} \rightarrow \text{закривання} \rightarrow \text{закривати} \rightarrow \text{проникнення} \rightarrow \text{проникати}.$$

кортеж $(x_1^0, x_2^0, \dots, x_{r_0}^0)$. З цього кортежу обирається певний елемент $-x_1(x_0)$, який у результаті застосування процедури A_1 стає заголовним словом словникової статті V_1 . Вибір елемента $x_1(x_0)$ не випадковий. Доцільно, щоб він перебував у певному гіперонімічному відношенні до x_0 (рід-вид, частина-ціле, комплекс-елемент тощо). Наступні кроки побудови такого ланцюга, тобто знаходження слів:

$$x_2(x_1), \dots, x_{n-1}(x_{n-2}), x_n(x_{n-1}),$$

які є результатами застосування операторів A_2, A_3, \dots, A_n , відповідно, є аналогічними. Зауважимо, що ми не вимагаємо, щоб на кожному кроці побудови ланцюга його елементи обирались тільки в одному, фіксованому відношенні до попередника, тобто не вимагаємо, щоб відношення між словом та його попередником у ланцюгові було сталим для всього ланцюга. Назвемо побудовані в такий спосіб ланцюги слів вигляду (2.49) *гіперланцюгами* Л-системи V . Гіперланцюги, для яких відношення між сусідніми словами є сталим упродовж всього гіперланцюга, називатимемо *однорідними*.

Розглянемо процес побудови гіперланцюгів на прикладі словникової статті із СУМ-20 із заголовним словом $x_0 =$ абажур. Його лексичне значення задається дефініцією «*Частина світильника, звичайно у вигляді ковпака, призначена для зосередження і відбиття світла та захисту очей від його впливу*», вихідні форми слів якої належать до кортежу $\langle V_0 \rangle$. За $x_1(x_0)$ виберемо слово *світильник*, яке, очевидно, перебуває зі словом *абажур* у відношенні *частина-ціле*. Одним зі значень цього слова є «*Прилад для освітлення – електричний або у вигляді лампи, площки, куди наливають олію і вставляють гніт*», з якого через дію автоморфізму A одержуємо: $A_2[x_1(x_0)] = x_2(x_1)$, де $x_2(x_1) =$ *прилад*. Слово *прилад* перебуває зі словом *світильник* у відношенні *рід-вид*. Продовжуючи цей процес, одержуємо такий гіперланцюг:

$$Z_1 = \text{абажур} \rightarrow \text{світильник} \rightarrow \text{прилад} \rightarrow \text{пристрій} \rightarrow \text{обладнання} \rightarrow \text{прилад}.$$

Як бачимо, на п'ятому кроці цього процесу гіперланцюг «зациквився», тобто знову повернувся до слова *прилад*, яке вперше з'явилося на третьому кроці. Такі гіперланцюги називатимемо *гіперциклами*. Для замкнених Л-систем, як впливає з теореми 2, будь-який достатньо довгий гіперланцюг є гіперциклом. Наведемо ще один приклад гіперциклу:

$$Z_2 = \text{абажур} \rightarrow \text{ковпак} \rightarrow \text{накривка} \rightarrow \text{закривання} \rightarrow \text{закривати} \rightarrow \text{проникнення} \rightarrow \text{проникати}.$$

Як бачимо, в цьому гіперциклі зустрічаються не лише іменники, але й дієслова. Отже, він є, так би мовити, двічі неоднорідним, оскільки він неоднорідний як у семантичному, так і граматичному відношеннях.

Гіперланцюги та гіперцикли певного слова, очевидно, становлять його семантичну характеристику. Це стає навіть формально очевидним при підстановці кожного наступного члена в попередній у формулі (2.50):

$$x_n(x_{n-1}(x_{n-2} \dots (x_2(x_1(x_0)) \dots)) \dots) = x_n(x_{n-1})$$

і розв'язуючи одержане рівняння відносно x_0 :

$$x_0 = f(x_1, x_2, \dots, x_{n-2}, x_{n-1}, x_n).$$

Вони можуть бути корисними при укладанні словника, адже кожний наступний член гіперланцюга або гіперциклу, будучи певним гіперонімом до свого попередника, мусить у своїй дефініції нести потенції до відповідної гіпонімії. Це означає, що лексикограф, укладаючи дефініцію лексичного значення, повинен стежити за тим, чи не може дефініційована реалія мати відповідні гіпоніми, і – в разі позитивної відповіді на це питання – забезпечувати в дефініції реалізацію такої можливості. Адже слово за самою своєю природою є носієм гіпо-гіперонімічних відношень, які повинні бути коректно відображені в лексикографічній системі тлумачного типу.

Слід зазначити, що процес побудови гіперланцюгів та гіперциклів, розглядуваний алгоритмічно, набуває рис певного логіко-лінгвістичного числення. Він і насправді подає приклад такого числення – ми називасмо його лексикографічним численням; більш докладно воно буде розглянуто нижче.

Описані конструкції мають ще один цікавий логіко-лінгвістичний аспект, пов'язаний з теоремою Геделя про неповноту [264].

Річ у тому, що замкненість гіперланцюгів та їх фактична редукованість до гіперциклів демонструє певну логічну хибу у дефініюванні (тавтологію), оскільки врешті-решт ситуація виглядає так, що слово x тлумачиться через слово y , а y через x . Розглядаючи словникові дефініції як аксіоми певної теорії та застосовуючи до такої формальної системи теорему Геделя про неповноту, доходимо висновку, що в такій теорії обов'язково мусять існувати твердження, які не можуть бути доведені або спростовані засобами самої теорії. Замкненість гіперланцюгів скінченної лексикографічної системи є маніфестацією її неповноти, за Геделем.

2.4. Лексикографічні середовища

У реальній дійсності об'єкти мови функціонують у їхній цілісності, не поділеній на окремі складники концептуального представлення. У лексикографічній системі це виявляється в тому, що при моделюванні мовних об'єктів засобами теорії Л-систем постає завдання інтеграції різних типів лексикографічних ефектів, а також поєднання й узгодження різнорідних (гетерогенних) лексикографічних структур. У свою чергу, це потребує узгодження між усіма елементами архітектури Л-систем, які підлягають процесу інтеграції.

Численні експерименти зі створення конкретних комп'ютерних реалізацій інтегрованих лінгвістичних об'єктів дозволили зробити висновок про необхідність побудови спеціального мовно-інформаційного середовища, яке б із самого початку було адаптованим до процесів інтегрування різних лексикографічних систем і містило необхідні засоби та конструктиви для здійснення зазначених процесів, а також фіксації їх результатів у вигляді інтегрованих лексикографічних систем, що мають відмінні лексикографічні структури. Як наслідок, було запропоновано поняття *лексикографічного середовища* [288; 224].

Визначення 10. Вважатимемо, що задано лексикографічне середовище (Л-середовище) ML , якщо:

1) задано клас $Ob\ ML$ елементів, кожен із яких є діаграмою вигляду (2.36) і представляє певну Л-систему (не обов'язково елементарну). Елементи з $Ob\ ML$ називаються об'єктами Л-середовища ML – позначатимемо їх великими латинськими літерами A, B, C, \dots ;

2) для кожної пари об'єктів A, B з ML задано множину $Hom_{ML}(A, B)$, яка називається множиною морфізмів A в B ; замість $f \in Hom_{ML}(A, B)$ також пишуть:

$f: A \rightarrow B$ або $A \xrightarrow{f} B$. При цьому $f: CM_A \rightarrow CM_B; f: INM_A \rightarrow INM_B; f: EXM_A \rightarrow EXM_B; f(\varphi_A) = \varphi_B; f(\psi_A) = \psi_B; f(\xi_A) = \xi_B$ та $f(\xi_A) \circ f(\psi_A) = f(\varphi_A)$;

3) для кожної трійки об'єктів (A, B, C) з ML задано відображення $\mu: Hom_{ML}(A, B) \times Hom_{ML}(B, C) \rightarrow Hom_{ML}(A, C)$ (образ $\mu(f, g)$ пари (f, g) , де $f \in Hom_{ML}(A, B)$, $g \in Hom_{ML}(B, C)$), буде позначатися $f \circ g$ або fg і називатися композицією морфізмів f і g ;

4) множини $Hom_{ML}(A, B)$ і композиція морфізмів задовольняють таким аксіомам:

(а) асоціативність: для кожної трійки морфізмів f, g, h :

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D \quad f(g h) = (f g)h;$$

(б) існування одиниці: для кожного $A \in \text{Ob } ML$ існує морфізм $1_A: A \rightarrow A$ ($1_A \in \text{Hom}_{ML}(A, A)$), такий що $1_A f = f$ і $g 1_A = g$ для довільних морфізмів $f \in \text{Hom}_{ML}(B, A)$ і $g \in \text{Hom}_{ML}(A, B)$;

(с) якщо пари (A, B) і (A', B') різні, перетин $\text{Hom}_{ML}(A, B)$ і $\text{Hom}_{ML}(A', B')$ порожній.

Нехай дано два лексикографічних середовища ML_1 і ML_2 . Коваріантний (відповідно контраваріантний) функтор F з ML_1 в ML_2 складається з:

(а) відображення $A \rightarrow F(A)$, яке зіставляє кожен об'єкт $A \in \text{Ob } ML_1$ з об'єктом $F(A) \in \text{Ob } ML_2$;

(б) відображень $F(A, B): \text{Hom}_{ML_1}(A, B) \rightarrow \text{Hom}_{ML_2}(F(A), F(B))$ – для коваріантного й $F(A, B): \text{Hom}_{ML_1}(A, B) \rightarrow \text{Hom}_{ML_2}(F(B), F(A))$ – для контраваріантного функторів, визначених для всіх пар (A, B) об'єктів з ML_1 і таких, що (якщо замість $F(A, B)(u)$ писати $F(u)$) $F(1_A) = 1_{F(A)}$ і $F(vu) = F(v)F(u)$ (відповідно $F(vu) = F(u)F(v)$).

2.5. Інтегровані Л-системи та методика їх побудови

Уведена в описаний вище спосіб структура Л-середовища є зручним формальним об'єктом для формування складних лексикографічних конструкцій, які поєднують в одне ціле багато окремих різнорідних (гетерогенних) Л-систем. Гетерогенність Л-систем, які підлягають інтеграції, є багатоаспектним поняттям. Ми розглядаємо Л-системи, неоднорідні за всіма рівнями архітектури – концептуальним, внутрішнім і зовнішнім. Такий підхід передбачає розроблення методів інтеграції концептуальних моделей, способів представлення даних та операційно-програмних платформ, а також узгодження зовнішніх представлень відповідних концептуальних схем та їх внутрішніх репрезентацій.

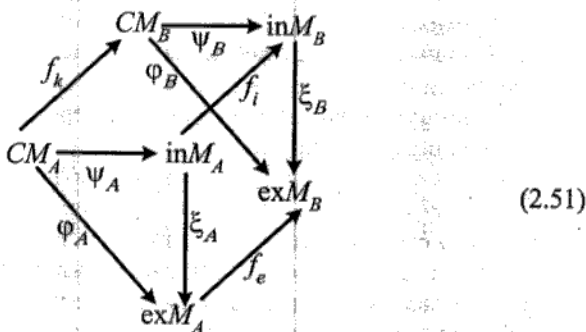
Під інтеграцією Л-систем як інформаційних систем розуміємо досягнення можливості одночасного та спільного використання прикладної програмою кількох інформаційних систем як єдиного цілого. Аналогічні визначення пропонує Державний стандарт України ДСТУ 2941-94: інтегрована система – сукупність двох або кількох взаємопов'язаних систем, у якій функціонування однієї з них залежить від результатів функціонування іншої (інших) так, що цю сукупність можна розглядати як єдину систему; інтегрування систем – об'єднання кількох систем різного призначення в єдину багатofункціональну систему.

Отже, для прикладної програми інтегрована сукупність різних лексикографічних баз даних (ЛБД) повинна мати вигляд як єдина ЛБД. Ці уяв-

лення певною мірою можуть бути перенесені й на традиційні словники, які також можуть являти собою інтегровані Л-системи, зокрема, тлумачний словник є одним із найяскравіших прикладів такої Л-системи.

У цьому підрозділі максимальний акцент зроблено на інтеграції Л-систем на концептуальному рівні. Базовими об'єктами, що підлягають інтеграції, є діаграми вигляду (2.36), які ми розглядаємо як об'єкти певного Л-середовища.

Нехай є два об'єкти A і B , а також $f: A \rightarrow B$, $f \in \text{Hom}_{ML}(A, B)$. Побудуємо f як спеціальний морфізм у вигляді трикомпонентного вектора: $f = (f_k, f_i, f_e)$, такого, що діаграма:



є комутативною – це еквівалентно справедливості таких рівностей:

$$\Psi_B \circ f_k = f_i \circ \Psi_A; \Phi_B \circ f_k = f_e \circ \Phi_A; \Xi_B \circ f_i = f_e \circ \Xi_A. \quad (2.52)$$

Діаграма (2.51) та рівності (2.52) надають необхідні формальні засоби, мовою яких можуть бути сформульовані процеси побудови інтеграційної архітектури, що впливає з наступного викладу.

Морфізм $f = (f_k, f_i, f_e), f: A \rightarrow B$ називається регулярним, якщо він задовольняє умовам:

- повної визначеності, згідно з якою будь-якому стану A відповідає один і тільки один стан B ;
- інтерпретованості, згідно з якою будь-якому елементу, який належить до множини $\{\sigma_A[\beta_A], RR\downarrow[V(A)]\}$, однозначно відповідає елемент із множини $\{\sigma_B[\beta_B], RR\downarrow[V(B)]\}$;
- відтворюваності, згідно з якою зміні будь-якого стану A , що здійснюється певним оператором з $\{\sigma_A[\beta_A], RR\downarrow[V(A)]\}$, відповідає адекватна зміна відповідного стану, котрий виконується певним оператором з $\{\sigma_B[\beta_B], RR\downarrow[V(B)]\}$.

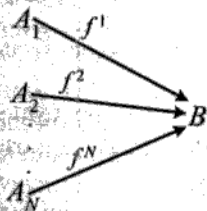
Поряд зі здатністю до інтеграції концептуально гетерогенних лексикографічних систем, що репрезентують різні мовні явища, не менш важ-

ливим аспектом інтеграційної архітектури є можливість досягнення високого ступеня незалежності прикладних програм від СКБД і забезпечення їх мобільності щодо СКБД різних типів. Проблема мобільності програм узагалі формулюється як забезпечення можливості виконання певної програми на різних комп'ютерних платформах без її зміни. Аналогічно визначається питання щодо мобільності програм стосовно СКБД.

Теоретично для цього існують такі можливості: 1) створення універсальної мови програмування й вимога її всезагального застосування; 2) забезпечення кожного комп'ютера компіляторами для всіх мов програмування при належній їх стандартизації; 3) введення платформи-незалежної проміжної мови з її інструментальною реалізацією на рівні віртуальної машини зі вбудованими інтерфейсами до будь-яких платформ; 4) застосування емуляційних методів; 5) використання комп'ютерних мереж за припущення, що хоча б один із мережевих комп'ютерів забезпечено необхідним компілятором.

Аналіз викладених пропозицій приводить до висновку про раціональність саме третьої з них. У різних варіаціях її намагаються реалізувати, наприклад, у системах типу JAVA, стандартизації машинно-незалежних мов SQL, SQL2 тощо. Зауважимо, зазначені пропозиції зовні аналогічні конструюванню спільної концептуальної моделі, що вкотре підтверджує положення про центральність саме концептуальної моделі в архітектурі інформаційної системи. Тому далі розглядатимемо саме розробку процесів інтеграції Л-систем на концептуальному рівні (або інтеграції концептуальних представлень Л-систем).

Зміст процедури інтеграції полягає ось у чому. Припустимо, що замість одного об'єкта A , наявного в діаграмі (2.51), маємо набір об'єктів A_1, A_2, \dots, A_N . Збудуємо вільноподібне відображення з морфізмами $f = (f^1, f^2, \dots, f^N)$:



(2.53)

$$f^p: A_p \rightarrow B, p = 1, 2, \dots, N,$$

де кожний f^p задається трикомпонентним вектором $f^p = (f_c^p, f_l^p, f_e^p)$ з властивостями (2.52).

У процесі застосування даних морфізмів використовується інтеграційна процедура, яка має такі етапи. Спочатку будується допоміжна Л-систе-

ма B з таких елементів: знакові системи $A(A_p)$; структури $\beta_p, \sigma_p[\beta_p]$ та процеси $RR_p \downarrow [V(\mathcal{P}(S))]$ всіх A_p . На цьому етапі відображення $f^p: A_p \rightarrow B$, $p=1, 2, \dots, N$ інтерпретується як відображення вкладення. Так отримуємо певну неелементарну Л-систему B з незалежними складниками A_p , $p=1, 2, \dots, N$.

Знакова система для Л-системи B будується як об'єднання: $A(B) = \cup A(A_p)$.

Далі вводимо спеціальну семантичну процедуру SEM , яка забезпечує ототожнення елементів структури, що збігаються бодай для двох систем із множини A_p , $p=1, 2, \dots, N$. Вважатимемо, що ця процедура здійснює ототожнення не лише імен семантично тотожних атрибутів, наявних у різних Л-системах A_p , $p=1, 2, \dots, N$, але й відповідних областей їхніх значень (доменів). При цьому можливі різні випадки перетинів структур, що належать до різних A_p , $p=1, 2, \dots, N$. Розглянемо їх докладніше.

Позначимо через $\varepsilon_{p_1 p_2 \dots p_k}[\beta]$ множини елементів структури (разом з їхніми доменами), які належать водночас до кожного з $A_{p_1}, A_{p_2}, \dots, A_{p_k}$, $1 \leq p_1 < p_2 < \dots < p_k \leq N$. Застосуємо до неї семантичну операцію:

$$SEM(\varepsilon_{p_1 p_2 \dots p_k}[\beta]) = \varepsilon_{p_1 p_2 \dots p_k}^{SEM}[\beta] \quad (2.54)$$

і в такий спосіб одержимо семантично тотожні спільні елементи структури Л-систем з номерами p_1, p_2, \dots, p_k . Наявність елементів (2.19) дає можливість побудови Л-системи зі структурою:

$$\sigma_{p_1 p_2 \dots p_k}^{SEM}[\beta] = (\varepsilon_{p_1 p_2 \dots p_k}^{SEM}[\beta], [\beta_{p_1}^{SEM}], [\beta_{p_2}^{SEM}], \dots, [\beta_{p_k}^{SEM}]; R^{SEM}), \quad (2.55)$$

де $[\beta_{p_i}^{SEM}]$, $i=p_1, p_2, \dots, p_k$ — елементи структури Л-систем A_i , в яких виконано ототожнення спільних елементів; R^{SEM} — об'єднання операцій, що діють у кожній Л-системі. Решта елементів структури Л-систем A_i , $i=p_1, p_2, \dots, p_k$, — позначимо їх через S_i , $i=p_1, p_2, \dots, p_k$ — залишається без зміни.

У такий спосіб Л-система B зі структурою:

$$\{\sigma_{p_1 p_2 \dots p_k}^{SEM}[\beta]; S_i, i=p_1, p_2, \dots, p_k\} \quad (2.56)$$

інтегрує Л-системи A_i , $i=p_1, p_2, \dots, p_k$.

Позначимо через $\varepsilon[\beta]$ множини елементів структури (разом з їх доменами, які належать до всіх A_p , $p=1, 2, \dots, N$):

$$\varepsilon[\beta] = \bigcap_p^N [\beta] A_p. \quad (2.57)$$

У результаті одержуємо Л-систему B зі знаковою системою $A(B) = \cup A(A_p)$ та структурою $\beta(B)$, $\sigma[\beta(B)]$, $RR \downarrow [V(B)]$, одержаною описаним вище спо-

собом. Особливо цікавим є випадок, коли певний елемент або певні елементи структури наявні в усіх L -системах, що підлягають інтеграції, тобто коли $\varepsilon[\beta] \neq \emptyset$. Тоді доцільно провести суцільне індексування всіх A_p за значеннями відповідних доменів, які належать до елементів $\varepsilon[\beta]$. Ці елементи структури набувають статусу вхідних до кожної з A_p . Як показує досвід, елементами структури відзначеного типу для природномовних систем виступають лексичні масиви (множини словоформ) усіх мов, що беруть участь в A_p . Отже, постає завдання побудови процедури природномовної індексації.

Побудована описаним способом L -система B називається інтеграцією L -систем $A_p, p = 1, 2, \dots, N$.

Структура лексикографічного середовища містить набір елементів, необхідний для представлення ефектів граматичної та лексичної семантики. Зокрема, це стосується і подання окремих відношень і мовних явищ, які абстрагуються з мовного континууму – словозміна, словотвір, орфоєпія, фразеологія, синонімія, антонімія тощо. Більше того, конструкція лексикографічного середовища допускає моделювання кожного із зазначених явищ та відношень окремо з наступною інтеграцією їх у єдиний лексикографічний комплекс.

Розділ 3

СИСТЕМО-ЛІНГВІСТИЧНІ АСПЕКТИ ТЛУМАЧНОГО СЛОВНИКА УКРАЇНСЬКОЇ МОВИ

3.1. Теорія Л-системи тлумачного Словника української мови

Перейдемо до викладу застосування теорії лексикографічних систем при побудові лексикографічного середовища тлумачного Словника української мови та розгляду системно-лінгвістичних ефектів, що імпліцитно містяться в структурі цього лексикографічного середовища. Застосування теорії Л-систем дозволило здійснити так званий парсинг (конверсію тексту даного словника в ЛБД) в автоматичному режимі для дуже складного лексикографічного об'єкта, яким є СУМ-11, – нам невідомі реальні приклади парсингу словників такого великого обсягу та складності – й на цій базі побудувати високо-ефективну комп'ютерну технологію укладання тлумачних словників.

Відзначимо ще один аспект СУМа – структурний, перше звернення до якого відбулося саме у наших працях [284; 287; 276]. Незважаючи на те, що про структуру словників взагалі та про структуру СУМа зокрема мовознавцями написано чимало, проте до наших праць з теорії лексикографічних систем зазначені структури розглядалися переважно як засіб до зручного та компактного представлення багатьох мовних фактів в одному об'єкті – словниковій статті. Тільки теорія лексикографічних систем надала концептуальний апарат та формальні засоби до розгляду лексикографічних структур як самостійних мовно-інформаційних об'єктів, що допускають абстрагування від конкретних текстових реалізацій у словникових статтях і, будучи представленими у вигляді певних формальних

Далі послуговуємося скороченням СУМ, якщо мова у викладі стосується як СУМ-11, так і СУМ-20. Якщо йтиметься про конкретний словник, то буде вказуватися, відповідно, СУМ-11 або СУМ-20.

об'єктів, набувають самостійної лінгвістичної інтерпретації та можуть застосовуватися для проведення цілком нових мовознавчих досліджень. Саме такі дослідження з використанням формальних репрезентантів лексикографічних структур СУМа було проведено в Українському мовно-інформаційному фонді, причому було одержано нові мовознавчі результати, досягти яких іншими методами, на нашу думку, практично неможливо. Саме використання апарату теорії лексикографічних систем дозволило відокремити лексикографічні структури від тексту словника і подивитися на них як на своєрідні маніфестанти мовної системи, в яких «зашифровано» певні закономірності, притаманні мові. Для обґрунтування викладеного звернімося до встановлення та аналізу лексикографічних структур СУМа.

Моделювання засобами теорії лексикографічних систем тлумачного Словника української мови є, мабуть, найсерйознішим випробуванням цієї теорії, оскільки СУМ є словником інтегрального типу, який містить велику кількість мовних феноменів, а його обсяг дозволяє вважати, що зазначені феномени подано у їх репрезентативному варіанті. Зауважимо, що побудова формальних лексикографічних структур СУМ-11, здійснена в УМІФі ще на початку 90-х років минулого століття, значною мірою зміцнила наше переконання у плідності концепції лексикографічних систем і спонукала до поглиблення її змісту у напрямі більшої строгості та прозорості основних засад теорії.

Перейдемо до викладу лексикографічної структури СУМа згідно з методологією теорії Л-систем.

У середині мовного потоку відбувається лексикографічний ефект спеціального типу, наслідком якого, згідно із загальною настановою теорії Л-систем, постає генерація такого комплексу елементарних інформаційних об'єктів, який кваліфікуємо як клас слів української мови $I^W(U)$. Відзначена кваліфікація, звичайно, не може слугувати визначенням цього комплексу, оскільки для цілей лексикографії потрібний більш досконалий опис лексичних одиниць, до визначення якого необхідно залучати лінгвістичні поняття, пов'язані з будовою слів, приписуванням їм значень певних граматичних категорій, походженням, поняттями значення та смислу, функціонуванням у контексті тощо. З цією метою в концептуальній моделі Л-системи визначаються такі види описів цього об'єкта:

$$\begin{array}{ll}
 V_f[I^W(U)] & \text{— фонетичний;} \\
 V_g[I^W(U)] & \text{— графічний;} \\
 V_m[I^W(U)] & \text{— морфемний;} \\
 V_{gr}[I^W(U)] & \text{— граматичний;} \\
 V_{sem}[I^W(U)] & \text{— семантичний;} \\
 V_{et}[I^W(U)] & \text{— етимологічний;} \\
 V_{sty}[I^W(U)] & \text{— стилістичний;} \\
 V_{st}[I^W(U)] & \text{— статистичний і т.д.}
 \end{array} \quad (3.1)$$

Кожен із цих описів або будь-яка їхня комбінація може формувати певну автономну Л-систему. Можлива також і побудова відображень цих Л-систем до певних стандартних моделей даних. Наприклад, морфемний опис $V_M[I^W(U)]$ у деякому наближенні можна подати в термінах реляційної моделі даних. Множину атрибутів (імен атрибутів) у цьому випадку можна визначити такою:

$$\begin{aligned} \Pi &: = \text{префікс}; K : = \text{корінь}; C : = \text{суфікс}; \\ M &: = \text{міжкоренева прокладка}; \Phi : = \text{флексія}. \end{aligned} \quad (3.2)$$

Відповідні домени, елементи доменів та назви цих інформаційних змінних позначимо через:

$$\begin{aligned} D(\Pi) &= \{\Pi_1, \Pi_2, \dots, \Pi_k\}, \\ D(K) &= \{K_1, K_2, \dots, K_p\}, \\ D(C) &= \{C_1, C_2, \dots, C_s\}, \\ D(M) &= \{M_1, M_2, \dots, M_m\}, \\ D(\Phi) &= \{\Phi_1, \Phi_2, \dots, \Phi_j\}. \end{aligned} \quad (3.3)$$

На декартовому добуткові доменів:

$$D(\Pi) \times D(K) \times D(C) \times D(M) \times D(\Phi)$$

визначається певна сукупність функцій (відношень концептуальної моделі):

$$\{r(\Pi, K, C, M, \Phi)\}, \quad (3.4)$$

множина значень якої ототожнюється з деякою власною підмножиною множини $I^W(U)$. Це, власне, і будуть реальні слова у мові з морфемним аналізом, алгоритми якого реалізовані набором відношень (3.4). Таким чином, відношення $r(\Pi, K, C, M, \Phi)$ (фактично, вони є правилами словотворення) задають закон композиції елементів з:

$$D(\Pi) \times D(K) \times D(C) \times D(M) \times D(\Phi)$$

так, що послідовність:

$$\Pi_\alpha * K_\beta * C_\gamma * M_\delta * \Phi_\epsilon \quad (3.5)$$

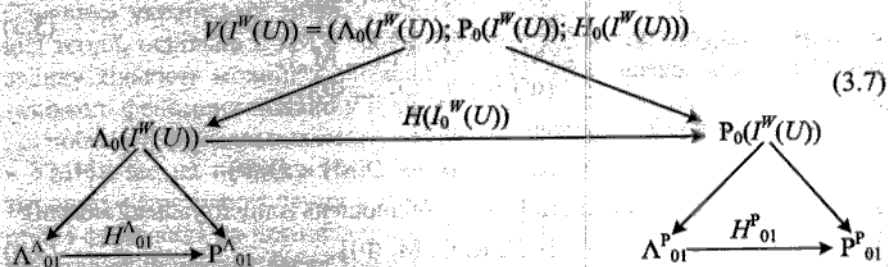
де зірочкою позначено конкатенацію, репрезентує реальне слово мови. Структура відповідної елементарної Л-системи тут визначається такою діаграмою:

$$\begin{array}{ccc} & V_M[I^W(U)] & \\ & \swarrow \quad \searrow & \\ \Lambda[I^W(U)] \equiv D(\Pi) \times D(K) \times D(C) \times D(M) \times D(\Phi) & \xrightarrow{H \equiv \{r(\Pi, K, C, M, \Phi)\}} & P[I^W(U)] \equiv \{\Pi_\alpha * K_\beta * C_\gamma * M_\delta * \Phi_\epsilon\}. \end{array} \quad (3.6)$$

Зауважимо, що виклад у даному випадку має спрощений характер і слугує виключно ілюстративним цілям.

Визначальною системоутвірною характеристикою комплексу ЕЮ «СЛОВО» виступає властивість значення, яка, у свою чергу, є носієм відношення «форма-зміст». Так, формальна частина ВФЗ, співвідносного зі значенням комплексу ЕЮ «СЛОВО», зосереджується навколо його властивостей у системі самої мови, у той час як змістова частина того самого відношення концентрується переважно навколо онтологічних вимірів цього комплексу, що стосуються буття як такого і виступають наслідком фундаментальної властивості слів репрезентувати реалії об'єктивної (позамовної) дійсності.

Зі сказаного випливає, що Л-система з класом ЕЮ «СЛОВО» повинна підтримувати структуру рекурсивної редукції як мінімум другого порядку і, отже, набувати вигляду:



На діаграмі (3.7) символом $I^W(U)$ позначено клас слів української мови; природною інтерпретацією структурного елемента $\Lambda_0(I^W(U))$ виступає його інтерпретація як носія граматичної семантики, а $P_0(I^W(U))$, відповідно, — лексичної семантики. Їх зв'язок і поєднання лінгвістичного об'єкта в єдине ціле забезпечує функція $H_0(I^W(U))$.

У свою чергу, структурний елемент $\Lambda_0(I^W(U))$, розглядуваний як репрезентант граматичної семантики, набуває інтерпретації як Л-система, що підтримує відношення словозміни, словотвору (та, можливо, певні інші морфолого-дериваційні відношення на рівні лексичної системи, а також, певною мірою, орфоепічні відношення — акцентуацію та вимову; для повного представлення орфоепічних явищ у цьому випадку слід застосувати рекурсивну редукцію третього порядку, ввівши в обіг відповідний лексикографічний ефект).

У конструкції репрезентанта лексичної семантики $P_0(I^W(U))$ представлено цілу низку семантичних відношень, причому явно виділено ієрархію лексичних значень (для кожної лексеми x вони зосереджені в структурному елементі $\Lambda_{01}^P(I^W(U))$ і представлені у вигляді формул тлумачень або словникових дефініцій) та відповідних мікроконтекстів (прикладів словживань), які репрезентовано в елементах $P_{01}^P(I^W(U))$. Відзначимо, що в

елементах $\Lambda_{01}^P(I^W(U))$ зосереджується і певна частина ідентифікаторів відношень синонімії, антонімії, меронімії, гіпо- та гіперонімії тощо.

Відповідно до загальної конструкції Л-системи в структурі тлумачної Л-системи виділяється множина $V(I^W(U))$ описів класу елементарних інформаційних одиниць – їхнім словниковим аналогом виступає множина словникових статей, а також ціла низка відношень як між певними підмножинами множини словникових статей, так і між окремими підмножинами структурних елементів словникових статей.

Деталізацію структури тлумачної Л-системи можна проводити різними способами:

1) з'ясувати, які типи лексикографічних ефектів належить подати в словнику, дослідити способи і розробити процедуру їх представлення у вигляді елементарних Л-систем, після чого застосувати до них процедуру інтеграції;

2) послідовно застосовувати процедуру рекурсивної редукції до елементарної Л-системи, представлені верхнім поверхом діаграми (3.7);

3) із самого початку побудувати структуру тлумачної Л-системи як елементарної Л-системи.

Найпослідовнішим і найкоректнішим вважаємо перший шлях, але водночас він є і найскладнішим. До нього близький і другий спосіб. У цьому ряді побудова елементарної тлумачної Л-системи видається найпростішим шляхом, тому скористаємося саме ним. Утім, це не виключає у подальшому застосування й процедури рекурсивної редукції, але вже до певних елементів структури, збудованої за третім способом тлумачної Л-системи.

Відповідно до загальної методології позначимо Л-систему СУМа символом $V(I_0^W(U))$, де клас $EIO I_0^W(U)$ є сукупністю українських слів у вихідній формі, а $V(I_0^W(U))$ являє собою множину їхніх описів:

$$V(I_0^W(U)) = \bigcup_{x \in I_0^W(U)} V(x), \quad (3.8)$$

де $V(x)$ інтерпретується як словникова стаття із заголовним словом x . Саме так відбувається центральність словникової статті в структурі тлумачних словників. Далі канонічним способом забезпечується подання основного відношення – ВФЗ:

$$\begin{array}{ccc} & V(I_0^W(U)) & \\ & \swarrow F \quad \searrow C & \\ \Lambda(I_0^W(U)) & \xrightarrow{H} & P(I_0^W(U)) \end{array} \quad (3.9)$$

$$F V(I_0^W(U)) = \Lambda(I_0^W(U)); \quad C V(I_0^W(U)) = P(I_0^W(U)); \quad \Lambda(I_0^W(U)) \cap P(I_0^W(U)) = \emptyset,$$

причому $H \circ F = C$, де символом « \circ » позначено композицію відображень;

$$\Lambda(I_0^W(U)) = \bigcup_{x \in I_0^W(U)} \Lambda(x); P(I_0^W(U)) = \bigcup_{x \in I_0^W(U)} P(x) \quad (3.10)$$

На $\Lambda(I_0^W(U))$ і $P(I_0^W(U))$ індукуються макроструктури:

$$F \sigma[\beta] = \lambda[\beta] \text{ і } C \sigma[\beta] = \rho[\beta] \quad (3.11)$$

й відповідні їм мікроструктури:

$$\lambda[\beta] \upharpoonright_{V(x)} \equiv \lambda(x); \rho[\beta] \upharpoonright_{V(x)} \equiv \rho(x) \quad (3.12)$$

як обмеження $\lambda[\beta]$ і $\rho[\beta]$ на $V(x)$. Саме в структурних елементах $\lambda(x)$ та $\rho(x)$ зосереджено максимальну лексико-граматичну та лексико-семантичну інформацію, тому побудова їх явного вигляду постає центральною проблемою глумачної лексикографії.

3.2. Прихована симетрія українського дієслова

Перейдемо до опису мовних ефектів, зосереджених у структурах $\lambda(x)$ та $\rho(x)$. Самі ці структури мають формальні репрезентанти, які абстрагуються з тексту СУМа, що дає можливість представити структуру останніх за допомогою буквально однієї формули. Повернімося до побудови формальних репрезентантів відзначених структур та аналізу тих лінгвістичних фактів, які з них випливають. Ми продемонструємо, що аналіз лексикографічних структур СУМа веде до встановлення певних фактів, які можна інтерпретувати як закони української мови в цілому, а не лише одного словника. Йдеться про певний закон, що має місце в організації морфологічної структури українського дієслова.

Цей, на наш погляд, повчальний приклад було помічено нами спочатку при дослідженні структури СУМ-11 [256; 291; 225], проте він виявився справедливим і для СУМ-20. У результаті нам вдалося встановити новий (як нам здається) тип лінгвістичної закономірності, названої нами *прихованою симетрією дієслова*.

Як ми неодноразово демонстрували, словник, на відміну від багатьох інших лінгвістичних об'єктів, являє собою добре структурований і достатньо формалізований матеріал. Тому, на перший погляд, може скластися враження, що знаходження в ньому якихось симетрій не є такою вже складною, та й такою цікавою задачею. Однак дослідження будови глумачного Словника української мови в 11 томах привело до відкриття такої його прихованої симетрії, яка представляє закон – причому, як ми переконаємося, досить строгий – уже для всієї системи української мови, а не лише для мовного матеріалу, використаного у згаданому словнику.

При вивченні структури СУМ-11 встановлено, що будова дієслівних словникових статей підпорядкована певній математичній закономірності, яку вдалося подати у вигляді цілком формальної структури. Виявилось, що існує формальний механізм, або закон, який «породжує» структури граматичної семантики українського дієслова (і, як буде показано далі, іменника). Незважаючи на те, що цьому закону підкоряються всі – без винятку – структури дієслівних словникових статей СУМ-11, він є настільки імпліцитним, що ніхто з дослідників, включаючи й авторів словника, не зміг його помітити.

Перейдемо до викладення суті цього явища. Відповідно до викладеної вище будови Л-системи СУМа в її структурі виділено множину реестрових (заголовних) слів: $W = \{X\}$, що слугують ідентифікаторами відповідних словникових статей $V(X)$. У структурі кожної словникової статті $V(X)$ виокремлюється ліва частина – $\Lambda(X)$, що відповідає за опис граматичної семантики реестрового слова X , і права частина – $P(X)$, де подається лексикографічне представлення лексичної семантики X . Крім цього, у лексикографічній системі СУМа визначається оператор $H: \Lambda(X) \rightarrow P(X)$, який забезпечує цілісність словникової статті і зв'язок між лексичною та граматичною семантикою (тобто між граматичною формою і лексичним значенням), що несе лексема X , а також цілу низку інших елементів (частково заданих імпліцитно), що відображають ті чи інші елементи лексикографічного опису лексичної системи.

Особливістю побудови СУМа є те, що в ньому, в одній словниковій статті, в лівій її частині, може фігурувати не тільки одне слово – реестрова одиниця, а декілька лексем, які утворюють реестровий ряд. Це відбувається тоді й тільки тоді, коли всі елементи реестрового ряду поєднані спільною лексичною семантикою, тобто вони мають один і той самий комплекс лексичних значень. Такий принцип побудови словникової статті є визначальним для СУМа.

Наприклад, у словниковій статті² з реестровою одиницею **ЗВИВАТИСЯ** ліва частина $\Lambda(X) =$ **ЗВИВАТИСЯ**, *рідко* **ІЗВИВАТИСЯ**, *аюся, аєтся, недок.* **ЗВІТИСЯ**, *зів'юся, зів'єшся, рідко* **ІЗВІТИСЯ**, *ізів'юся, ізів'єшся, діал.* **ЗВІНУТИСЯ** і *рідко* **ІЗВІНУТИСЯ**, *нуся, нешся, док.* має реестровий ряд, що складається з шести членів: **ЗВИВАТИСЯ**; **ІЗВИВАТИСЯ**; **ЗВІТИСЯ**; **ІЗВІТИСЯ**; **ЗВІНУТИСЯ**; **ІЗВІНУТИСЯ**, поєднаних спільною для всіх членів цього ряду лексичною семантикою, яка виражається спільним для них комплексом лексичних значень:

² Приклади словникових статей у цьому підрозділі відповідають СУМ-20.

- $P(X) = 1$. Скручуватися, закручуватися в трубку, кільце і т. ін.
2. З'єднуючись, спітатися до купи, разом.
 3. Вигинатися, нахилиючись у різні боки // Борсатися, корчитися (перев. від болю) // *тільки недок.* Згинатися, улесливо кланяючись // *тільки 3 ос.* Маги вигини, зигзаги, бути розташованим по звивистій лінії.
 4. *тільки недок., перен.* Хитрувати, викручуватися, пристосовуючись до обставин.
 5. Підніматися вгору клубами, вихором (про дим, пару, куряву і т. ін.).
 6. Стрімко злітати вгору (про птахів, літальні апарати і т. ін.) // Спрямуватися вгору // Підніматися, ставати на диби (про коня) // Гучно, високо звучати (про голос, музику).
 7. *перев. недок.* Рухатися в різних напрямках, крутитися, метушитися // Швидко повертатися, рухатися, справляючись з якою-небудь роботою // Літати, кружляти.
 8. *тільки недок.* Розвіватися, маяти в повітрі (про прапор, корогву, хустку, плащ і т. ін.) // Спадати хвилясто.
 9. *тільки недок., перен.* Підлесливо крутитися, упадати біля кого-небудь, домагаючись уваги, ласки.
 10. *тільки недок.* Пас. до **звива́ти**.

При цьому два перших члени реестрового ряду (**ЗВИВА́ТИСЯ; ІЗВИВА́ТИСЯ**) належать до недоконаного виду, а решта чотири (**ЗВІТИСЯ; ІЗВІТИСЯ; ЗВІНУТИСЯ; ІЗВІНУТИСЯ**) – до деконаного.

Така особливість представлення словникових статей у СУМі в сполученні з морфологічною будовою української мови веде до прихованої симетрії в їх структурі й підкоряється певним закономірностям. При їх установленні використовуватимемо методологію дослідження, яка формулюється у вигляді наступних вимог до теорії, яка мусить:

- 1) базуватися на мінімумі аксіом (постулатів). Припущення *ad hoc* вважаються неприпустимими;
- 2) задовільно пояснювати усі явища, що належать до досліджуваного класу;
- 3) мати певну передбачувальну силу і правильно прогнозувати нові явища і факти.

Ми свідомі того, що навіть у таких, на перший погляд, строгих вимогах до теорії є якась невимовна неточність, яка полягає в іманентній нечіткості, притаманній процесу побудови словникових статей і визначенні комплексів лексичних значень. Адже авторами-укладачами словникових статей є лексикографи, не вільні від певного (а точніше – непевного) суб'єктивізму, та й у цілому колектив лексикографів (а великі словники,

завичай, укладаються й значними колективами науковців) являє собою те, що можна охарактеризувати як колективний суб'єкт. Отже, семантичні комплекси, які представляють лексичну семантику реєстрових рядів, слід розглядати за модулем суб'єктивності лексикографічного колективу. Проте навіть у рамках такої суб'єктивності виявляється, що будова дієслівних словникових статей СУМа підкоряється певним прихованим і неусвідомлюваним навіть його авторами закономірностям.

А саме: якщо розглядати тексти $\Lambda(X)$, де X пробігає множину дієслів української мови, абстрагуючись від їхнього змісту, як лінійні ланцюжки символів, то можна помітити, що в структурі цих ланцюжків виділяються певні підланцюжки. Зазначені виділення мають характер вкладень, визначених інваріантним способом так, що різні підланцюжки не перетинаються один із одним.

Визначена закономірність будови ланцюжків $\Lambda(X)$ є такою регулярною, що дозволяє сформулювати її аксіоматично у вигляді трьох постулатів.

Перший постулат є цілком очевидним:

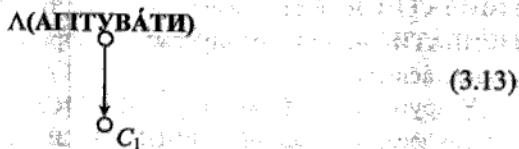
(1) Кожне українське дієслово реалізується в мові лексемою (із визначеною і зафіксованою лексичною семантикою), якій притаманні одне або два значення категорії виду.

Лексикографічне представлення належності лексеми до певного значення атрибута «ВИД» (як воно подано в СУМі) будемо називати *видовим комплексом*, або просто *комплексом* і позначати символами C_1 та C_2 .

Наприклад, у словниковій статті з $\Lambda(X) = \text{АГІТУВАТИ}$, *ую, уещ, недок.*, є один видовий комплекс: $C_1 = \text{АГІТУВАТИ}$, *ую, уещ*.

Таким чином, тут маємо вкладення першого рівня: $C_1 \subset \Lambda(X) = \Lambda(\text{АГІТУВАТИ})$.

Цей факт подано в графічній формі:

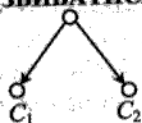


де відношення вкладення « \subset » представлено стрілкою.

У наведеній нами вище словниковій статті $\Lambda(\text{ЗВИВА́ТИСЯ})$ наявні два видових комплекси: $C_1 = \text{ЗВИВА́ТИСЯ}$, *рідко ІЗВИВА́ТИСЯ*, *аюся, а́сся та* $C_2 = \text{ЗВИ́ТИСЯ}$, *зів'юся, зів'єся, рідко ІЗВИ́ТИСЯ*, *зів'юся, зів'єся, діал. ЗВИ́НУТИСЯ і рідко ІЗВИ́НУТИСЯ*, *нуся, неся*. При цьому

комплекс C_1 стосується недоконаного виду, а комплекс C_2 – доконаного. Цей лінгвістичний факт, очевидно, має таке графічне представлення:

$\Lambda(\text{ЗВИВА́ТИСЯ})$



(3.14)

Другий постулат:

(2) Кожна дієслівна лексема з визначеною і зафіксованою лексичною семантикою та конкретним значенням виду може реалізуватися дієсловами, що належать не більше як до трьох різних словозмінних (парадигматичних) класів.

Ідентифікаторами словозмінних класів у СУМі є парадигматичні показники, якими, як правило, виступають набори квазіфлексій першої та другої особи однини теперішнього часу; якщо їх виявляється недостатньо для визначення системи словозміни, то наводяться деякі інші форми (про це докладно див. підрозділ 5.4. Б. Дієслова). Так чи інакше, але набір парадигматичних показників обирається достатнім для ідентифікації словозмінної парадигми конкретної лексики.

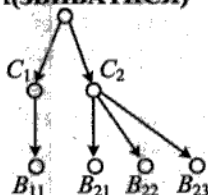
Для лексикографічного представлення цих фактів вживатимемо термін *парадигматичний блок*, або просто *блок*, який позначено символами B_{ij} , де індекс «i» відповідає номеру видового комплексу, до якого належить цей блок, і набуває значення 1, 2, а індекс «j» нумерує блоки всередині видового комплексу і набуває значення від 1 до 3.

Розглянемо для прикладу блокову структуру словникової статті $\Lambda(\text{ЗВИВА́ТИСЯ})$. У першому комплексі: $C_1 = \text{ЗВИВА́ТИСЯ}$, *рідко* ІЗВИВА́ТИСЯ , *аюся*, *аєшся*, є один блок: $B_{11} = \text{ЗВИВА́ТИСЯ}$, *рідко* ІЗВИВА́ТИСЯ , що поєднується набором парадигматичних показників «*аюся, аєшся*».

У другому комплексі: $C_2 = \text{ЗВІ́ТИСЯ}$, *зів'юся, зів'єшся, рідко* ІЗВІ́ТИСЯ , *ізів'юся, ізів'єшся, діал.* ЗВІ́НУТИСЯ і *рідко* ІЗВІ́НУТИСЯ , *нуса, нешся*, максимально можливе число блоків – три, а саме: $B_{21} = \text{ЗВІ́ТИСЯ}$, який визначається показниками «*зів'юся, зів'єшся*»; $B_{22} = \text{рідко } \text{ІЗВІ́ТИСЯ}$, який визначається показниками «*ізів'юся, ізів'єшся*» та $B_{23} = \text{діал. } \text{ЗВІ́НУТИСЯ}$ і *рідко* ІЗВІ́НУТИСЯ , який визначається показниками «*нуса, нешся*».

Графічне представлення блокової структури цієї словникової статті:

$\Lambda(\text{ЗВІВА́ТИСЯ})$



(3.15)

Третій постулат:

(3) Кожна дієслівна лексема з конкретним значенням виду, визначеною належністю до конкретного словозмінного класу і зафіксованою лексичною семантикою може реалізовуватися дієсловами, які мають не більше чотирьох фонетичних варіантів (як правило, це префіксальна та коренева варіації, але трапляється і суфіксальна).

Для лексикографічного представлення цього факту вживатиметься термін *компонент*. Компоненти позначаються символами K_{ij} , де індекс «i» відповідає номеру видового комплексу, до якого належить блок з номером «j», а індекс «r» пробігає значення від 1 до 4 залежно від того, скільки компонентів має конкретний блок.

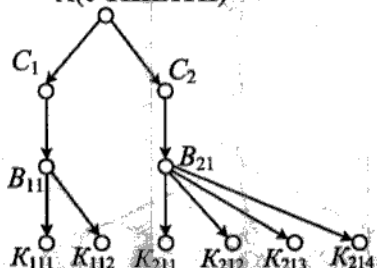
Для ілюстрації ефекту компонентності розглянемо словникову статтю:

$\Lambda(X) = \text{УТИНА́ТИ (ВТИНА́ТИ)}$, аю, аеш, *недок.*, УТНУ́ТИ (ВТНУ́ТИ) , УТЯ́ТИ (ВТЯ́ТИ) , утну́, утнш, *док.*, *кого, що і без дод.*

Тут, очевидно, є два видових комплекси. У першому комплексі $C_1 = \text{УТИНА́ТИ (ВТИНА́ТИ)}$, аю, аеш є лише один блок: $B_{11} = \text{УТИНА́ТИ (ВТИНА́ТИ)}$, до якого, у свою чергу, належать два компоненти: $K_{111} = \text{УТИНА́ТИ}$; $K_{112} = \text{ВТИНА́ТИ}$. У другому комплексі: $C_2 = \text{УТНУ́ТИ (ВТНУ́ТИ)}$, УТЯ́ТИ (ВТЯ́ТИ) , утну́, утнш є також лише один блок: $B_{21} = \text{УТНУ́ТИ (ВТНУ́ТИ)}$, УТЯ́ТИ (ВТЯ́ТИ) . Але в цьому блоці наявне максимальне число компонентів – чотири, а саме: $K_{211} = \text{УТНУ́ТИ}$; $K_{212} = \text{ВТНУ́ТИ}$; $K_{213} = \text{УТЯ́ТИ}$; $K_{214} = \text{ВТЯ́ТИ}$.

Графічно цей факт має таке представлення:

$\Lambda(\text{УТИНА́ТИ})$



(3.16)

З описаних властивостей композиції дієслівних $\Lambda(X)$ випливає висновок про існування особливої лінгвістичної закономірності в їх структурі, яка формально подається системою вкладень:

$$\Lambda(X) \supset C \supset B \supset K, \quad (3.17)$$

де символом C позначено сукупність видових комплексів, що належать до даного $\Lambda(X)$ (не більше двох комплексів); символом B – сукупність парадигматичних блоків (не більше трьох блоків у кожному комплексі); K – сукупність компонентів (не більше чотирьох компонентів у кожному блоці).

Назвемо відзначену закономірність правилом «1–2–3–4». Цікаво, що в цьому правилі поєднуються різні за своїм характером граматичні, структурно-морфологічні та лексико-семантичні явища, які задають закон заповнення структурних елементів лівих частин дієслівних словникових статей СУМа. Правило «1–2–3–4» насправді подає певну – приховану – симетрію в будові $\Lambda(X)$, допускаючи лексикографічні структури лише цілком визначених типів. Отже, не випадково, що відзначений факт пройшов непоміченим навіть для лексикографів – авторів концепції і структури СУМ-11.

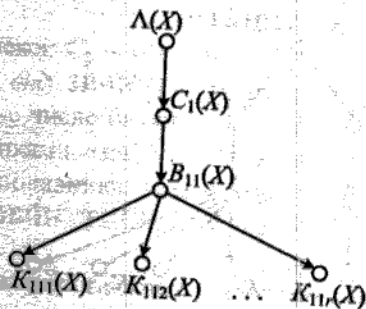
Оскільки будова дієслівних $\Lambda(X)$ має певну симетрію, то зрозуміло, що мусить існувати і спосіб її формалізації, на підставі якого можливе вичерпне перерахування допустимих цією симетрією лексикографічних типів $\Lambda(X)$.

Для встановлення і формалізації цієї симетрії прийнято позначення. Символом I^i позначено множину чисел 1, 2, 3, 4:

$$i \in I^i, i = 1, 2, 3, 4, \quad (3.18)$$

елементами якої нумеруються дієслова-компоненти, визначені третім постулатом – тобто фонетичні варіанти дієслів, яким властиве конкретне значення словозмінних параметрів.

Якщо ж застосувати разом усі три аксіоми, то будь-який з елементів множини (3.18) визначатиме певний видовий комплекс у такий спосіб. Фіксація числа r ($r = 1, 2, 3, 4$) сигналізує, що маємо дієслова, які характеризуються одним значенням атрибута «ВИД» із притаманним їм одним набором значень парадигматичних показників, а кількість фонетичних варіантів дорівнює r . Тоді, якщо відношення вкладення позначити стрілкою, тобто: $A \supseteq B \Rightarrow A \rightarrow B$, цей факт можна зобразити графом:



(3.19)

де символ $\Lambda(X)$ – ліва частина словникової статті з реєстровою одиницею X ; $C_1(X)$ – її єдиний комплекс; $B_{11}(X)$ – єдиний блок цього комплексу та $K_{111}(X), K_{112}(X), \dots, K_{11r}(X)$ – відповідно 1-а, 2-а, ..., r -а компоненти зазначеного блоку, де r може набувати значення від 1 до 4.

Очевидно, що декартів добуток $I^1 \times I^1 \equiv I^2$ параметризує структури двоблокових комплексів, а $I^1 \times I^1 \times I^1 \equiv I^3$ – структури триблокових комплексів. Взагалі одержуємо таку картину:

Таблиця 3.1. Параметризація структур видових комплексів

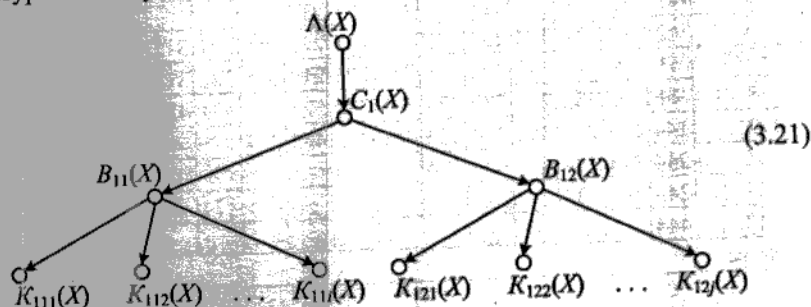
Одноблокові (I^1)						Двоблокові (I^2)									
1						11	12	13	14						
2						21	22	23	24						
3						31	32	33	34						
4						41	42	43	44						
Триблокові (I^3)															
111	112	113	114	211	212	213	214	311	312	313	314	411	412	413	414
121	122	123	124	221	222	223	224	321	322	323	324	421	422	423	424
131	132	133	134	231	232	233	234	331	332	333	334	431	432	433	434
141	142	143	144	241	242	243	244	341	342	343	344	441	442	443	444

У такий спосіб структура будь-якого видового комплексу представлена однією з трьох сигнатур:

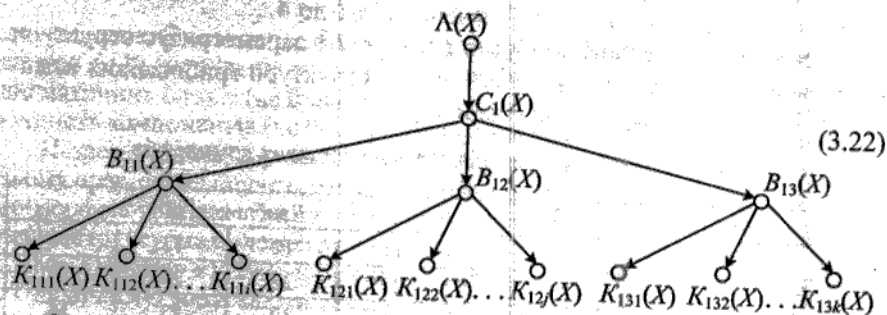
$$(i), (ij), (ijk), \quad (3.20)$$

у кожній з яких числа i, j, k незалежно пробігають множину $I^1 = \{1, 2, 3, 4\}$. Сигнатура (i) відповідає граф типу (3.19).

Сигнатура (ij) представляє комплекс, який має два парадигматичних блоки, у першому з яких наявні i , а в другому – j компоненти. Граф цього сигнатурного типу виглядає так:



Сигнатура $(ij k)$ представляє комплекс, який має три парадигматичних блоки, у першому з яких наявні i , у другому $-j$, а у третьому $-k$ компоненти. Графічно це зображається так:



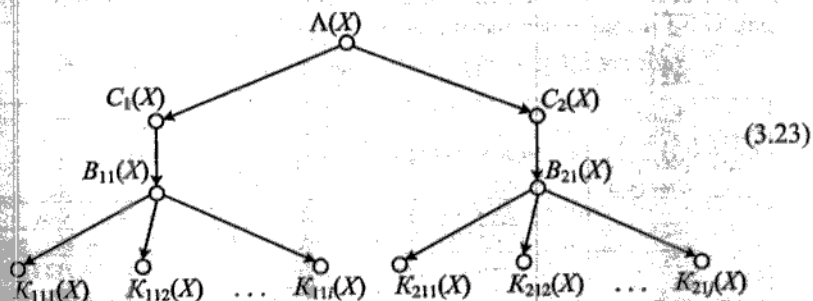
Отже, одноблокових сигнатур може бути максимум 4 (1×4), двоблокових – 16 (4×4), триблокових – 64 ($4 \times 4 \times 4$). Таким чином, видовий комплекс може реалізовуватися максимум 84 різними сигнатурами ($4 + 16 + 64 = 84$).

Загальний розподіл сигнатур за кількістю елементів реєстрового ряду наведено в табл. 3.2.

Таблиця 3.2. Розподіл сигнатур за кількістю одиниць реєстрового ряду

Кількість слів у реєстровому ряду комплексу	Сигнатури, через які реалізуються словникові статті з певним числом членів реєстрового ряду														Число сигнатур		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14			
1	1																1
2	2	11															2
3	3	21	12	111													4
4	4	31	22	13	121	112	211										7
5	41	32	23	14	131	122	113	221	212	311							10
6	42	33	24	141	132	123	114	231	222	213	321	312	411				13
7	43	34	142	133	124	241	232	223	214	331	322	313	421	412			14
8	44	143	134	242	233	224	341	332	323	314	431	422	413				13
9	144	243	234	342	333	324	441	432	423	414							10
10	244	343	334	442	433	424											6
11	244	443	434														3
12	444																1
Разом																	84

Словникова стаття, яка містить обидва видових комплекси, зображається парою (α, β) , де α і β незалежно пробігають множини сигнатур, поданих у формулі (3.20). Зобразимо це графом:



Очевидно, що сигнатур типу (α, β) може бути щонайбільше 7056 (84×84), а отже, загальне число структур, які задовольняють аксіоми (1)–(3), теоретично може дорівнювати 7140 ($84 + 7056$).

Таким чином, прихована симетрія дієслівних $\Lambda(X)$ приводить до того, що необмежена – на перший погляд – кількість структур словникових статей редукується до скінченного, точно визначеного цілого числа.

Побудований формалізм, структура сигнатур та їх розподіл є прямими наслідками постулатів (1)–(3). Самі ж ці постулати сформульовані як результат спостережень над текстами дієслівних словникових статей СУМ-11 на початку 90-х років, коли базу даних словника ще не було створено. Тоді ж постало питання: чи всі дієслівні словникові статті СУМ-11 задовольняють ці постулати і чи не існує винятків із правила «1-2-3-4»?

Відповідь на це питання в той час одержати було дуже складно, оскільки продивитися в «ручному» режимі структури понад 41 тис. дієслівних словникових статей, розроблених у СУМ-11, нереально. Це завдання ми розв'язали лише у 2001 році на основі сформованої лексикографічної бази даних СУМ-11, що дало можливість створити інструментальний комплекс для його модернізації та проведення на ньому ряду досліджень. Зокрема були вивчені й структури дієслівних $\Lambda(X)$ стосовно їхньої відповідності постулатам (1)–(3) та правилу «1-2-3-4». В результаті проведеного обчислювального експерименту над базою даних СУМ-11 з

потенційно можливих 7140 класів дієслівних $\Lambda(X)$ у СУМі реально виявлено всього 52 класи³.

У табл. 3.3 наведено узагальнені результати розподілу дієслів сучасної української мови за 52 сигнатурами. Слід наголосити, що ця класифікація не дає хиб і на масиві нового 20-томного СУМа.

Таблиця 3.3. Структурні класи дієслів, реалізовані у СУМі

№	Сигнатура	Кількість слів у реєстровому ряді	Структурна формула	Приклад
1.	1	1	$\Lambda(X) = C_1(B_{11}(K_{11}))$	АГТУВАТИ , <i>у́ю, у́еш, недок.</i>
2.	2	2	$\Lambda(X) = C_1(B_{11}(K_{11} + K_{112}))$	ЗАРЕВІТИ, ЗАРЕВІТИ , <i>в́у, в́еш, док.</i>
3.	3	3	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112} + K_{113}))$	ЗІСТАРІТИСЯ, ЗОСТАРІТИСЯ, ЗІСТАРТИСЯ , <i>і́юєш, і́єшш, док., розм.</i>
4.	4	4	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112} + K_{113} + K_{114}))$	УДЕРТИ (ВДЕРТИ), УДРАТИ (ВДРАТИ) , <i>удеру́, удереш, док., що і рідко без прям. дод., розм.</i>
5.	11	2	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121}))$	ДВИГНУТИ , <i>ну́, не́ш, ДВІЙНУТИ</i> , <i>ну, двиш, док., розм.</i>
6.	12	3	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121} + K_{122}))$	ОДУБИТИ , <i>і́ю, і́єш, рідко ОДУБНУТИ, ОДУБИТИ</i> , <i>бну, бнеш; мин. ч. одубів, рідко одубнув і одуб, ла, ло; док., розм.</i>
7.	21	3	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112}) + B_{12}(K_{121}))$	НЕДОГЛЯДИТИ, НЕДОГЛЯДТИ , <i>джу, дши, НЕДОГЛЯНУТИ</i> , <i>ну, неш, док., кого, чого, що, за ким – чим і без дод.</i>
8.	12	4	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112}) + B_{12}(K_{121} + K_{122}))$	ВБУХАТИ (УБУХАТИ) , <i>аю, аеш, ВБУХНУТИ (УБУХНУТИ)</i> , <i>ну, неш, док., що, фам.</i>
9.	111	3	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121}) + B_{13}(K_{131}))$	ПОНАЇЗЖАТИ , <i>а́ють, ПОНАЇЗДИТИ</i> , <i>діть, ПОНАЇХАТИ</i> , <i>а́ють, док.</i>
10.	112	4	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121}) + B_{13}(K_{131}) + K_{132})$	КМІТИТИ , <i>кмічу́, кмітиш, рідко КМЕТІТИ</i> , <i>кметчу́, кметитиш, КМІТУВАТИ, рідко КМЕТУВАТИ</i> , <i>у́ю, у́еш, недок.</i>
11.	1.1	2	$\Lambda(X) = C_1(B_{11}(K_{111})) + C_2(B_{21}(K_{211}))$	ВИГОВОРІЮВАТИ , <i>юю, юєш, недок., ВИГОВОРІТИ</i> , <i>рю, риш, док.</i>
12.	1.2	3	$\Lambda(X) = C_1(B_{11}(K_{111})) + C_2(B_{21}(K_{211} + K_{212}))$	НАДЯГАТИ , <i>а́ю, а́еш, недок., НАДЯГНУТИ, НАДЯГТИ</i> , <i>дягну́, дягнеш, док., що.</i>
13.	1.3	4	$\Lambda(X) = C_1(B_{11}(K_{111})) + C_2(B_{21}(K_{211} + K_{212} + K_{213}))$	ОБПІНАТИ , <i>а́ю, а́єш, недок., ОБП'ЯСТИ, ОБП'ЯСТИ</i> , <i>обпну́ти, обпну́, обпнеш, док., кого, що, розм.</i>

³ Докладно цей експеримент описано в монографії «Дієслово в лексикографічній системі» авторів Рабульця О. Г., Сухарини Н.М., Широкова В.А., Якименка К.М. [225]

№	Сигна-гура	Кількість слів у реєстровому ряді	Структурна формула	Приклад
14.	I.4	5	$\Lambda(X) = C_1 (B_{11}(K_{111})) + C_2 (B_{21}(K_{211} + K_{212} + K_{213} + K_{214}))$	СПИНАТИСЯ , аюся, аєшся, <i>недок.</i> , СПР'ЯСТИСЯ , ЗПР'ЯСТИСЯ , ЗПНУТИСЯ , ЗП'ЯТИСЯ , <i>зпну́ся, зпнє́шся, док.</i>
15.	I.11	3	$\Lambda(X) = C_1 (B_{11}(K_{111})) + C_2 (B_{21}(K_{211}) + B_{22}(K_{221}))$	НАВИВАТИ ¹ , аю, аєш, <i>недок.</i> , НАВИТИ , в'ю, в'єш, НАВИНУТИ , ну, нєш, <i>док., що.</i>
16.	I.12	4	$\Lambda(X) = C_1 (B_{11}(K_{111})) + C_2 (B_{21}(K_{211}) + B_{22}(K_{221} + K_{222}))$	ДОГЛЯДАТИ , аю, аєш, <i>недок.</i> , ДОГЛЯНУТИ , ну, нєш, <i>розм.</i> ДОГЛЯДТИ і ДОГЛЕДТИ , джу, диш, <i>док.</i>
17.	I.2	4	$\Lambda(X) = C_1 (B_{11}(K_{111})) + C_2 (B_{21}(K_{211} + K_{212}) + B_{22}(K_{221}))$	РОЗТАВАТИ , аю, аєш, <i>недок.</i> , РОЗТАНУТИ , РОЗТАТИ , аю, аєш, <i>рідко</i> РОЗТАЯТИ , аю, аєш, <i>док.</i>
18.	I.111	4	$\Lambda(X) = C_1 (B_{11}(K_{111})) + C_2 (B_{21}(K_{211}) + B_{22}(K_{221}) + B_{23}(K_{231}))$	НАСАДЖУВАТИ ¹ , аю, уєш, <i>недок.</i> , НАСАДИТИ , саджу, садиш, НАСАДЖАТИ , аю, аєш, <i>розм.</i> НАСАДОВИТИ , довлію, довиш; <i>мн.</i> <i>Насадовлять; док.</i>
19.	11.1	3	$\Lambda(X) = C_1 (B_{11}(K_{111}) + B_{12}(K_{121})) + C_2 (B_{21}(K_{211}))$	ОБСИПАТИ , аю, аєш, <i>рідко</i> ОБСИПУВАТИ , аю, уєш, <i>недок.</i> ОБСИПАТИ , плю, плєш; <i>мн.</i> <i>обсиплють; док.</i>
20.	11.2	4	$\Lambda(X) = C_1 (B_{11}(K_{111}) + B_{12}(K_{121})) + C_2 (B_{21}(K_{211} + K_{212}))$	НАДИХАТИ , аю, аєш, НАДИХУВАТИ , аю, уєш, <i>недок.</i> , НАТХНУТИ , НАДИХНУТИ , ну, нєш, <i>док., кого і без дод.</i>
21.	12.1	4	$\Lambda(X) = C_1 (B_{11}(K_{111}) + B_{12}(K_{121} + K_{122})) + C_2 (B_{21}(K_{211}))$	ВИПОРСКУВАТИ , аю, уєш, ВИПОРСАТИ , <i>рідко</i> ВИПОРКАТИ , аю, аєш, <i>недок.</i> , ВИПОРСНУТИ , ну, нєш, <i>док.</i>
22.	12.2	5	$\Lambda(X) = C_1 (B_{11}(K_{111}) + B_{12}(K_{121} + K_{122})) + C_2 (B_{21}(K_{211} + K_{212}))$	РОЗКОРИНЮВАТИСЯ , юється, РОЗКОРЕНЯТИСЯ , РОЗКОРИНЯТИСЯ , яється, <i>недок.</i> , РОЗКОРЕНИТИСЯ , РОЗКОРИНИТИСЯ , яється, <i>док.</i>
23.	11.11	4	$\Lambda(X) = C_1 (B_{11}(K_{111}) + B_{12}(K_{121})) + C_2 (B_{21}(K_{211}) + B_{22}(K_{221}))$	ЗАВЧАТИ , аю, аєш, <i>рідше</i> ЗАУЧУВАТИ , аю, уєш, <i>недок.</i> , ЗАВЧИТИ , чу, чєш, <i>рідше</i> ЗАУЧИТИ , учу, учєш, <i>док., що.</i>
24.	11.12	5	$\Lambda(X) = C_1 (B_{11}(K_{111}) + B_{12}(K_{121})) + C_2 (B_{21}(K_{211}) + B_{22}(K_{221} + K_{222}))$	ВІДДИХАТИ , аю, аєш, <i>рідко</i> ВІДДИХУВАТИ , аю, уєш, <i>недок.</i> , ВІДДИХАТИ , віддіхаю, віддіхаєш і віддішу, віддішєш, ВІДДИХНУТИ , ВІДТХНУТИ , ну, нєш, <i>док.</i>

№	Сигна-тура	Кількість слів у реєстровому ряді	Структурна формула	Приклад
25.	11.21	5	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121})) + C_2(B_{21}(K_{211} + K_{212}) + B_{22}(K_{221}))$	РОЗГЛЯДАТИ , аю, аеш, <i>рідко</i> РОЗГЛЯДУВАТИ / РОЗГЛЄДЖУВАТИ , ую, уеш, <i>недок.</i> , РОЗГЛЯДТИ , РОЗГЛЄДТИ , джу, дши, РОЗГЛЯНУТИ , ну, неш; <i>мин. ч.</i> розглядів, розгледів, діла, ло і розглянув, нула, до; <i>док.</i>
26.	12.11	5	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{21}(K_{211} + K_{212})) + C_2(B_{21}(K_{211}) + B_{22}(K_{221}))$	РОЗСТІБАТИ , аю, аеш, <i>рідше</i> РОЗСТІБУВАТИ / РОЗСТЬОБУВАТИ , ую, уеш, <i>недок.</i> , РОЗСТЕБНУТИ , стебнї, стебнеш, РОЗСТІБНУТИ , стібнї, стібнеш; <i>док., що.</i>
27.	11.111	5	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{21}(K_{211})) + C_2(B_{21}(K_{211}) + B_{22}(K_{221}) + B_{23}(K_{231}))$	РОЗЛІНІВОВАТИ , іюво, іюеш, РОЗЛІНОВУВАТИ , ую, уеш, <i>недок.</i> , РОЗЛІНОВАТИ , іюво, іюеш, РОЗЛІНЯТИ , ію, іеш, РОЗЛІНУВАТИ , ію, іеш, <i>док., що.</i>
28.	111.1	4	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121}) + B_{13}(K_{131})) + C_2(B_{21}(K_{211}))$	ВИКУПОВАТИ , ую, уеш, <i>розм.</i> ВИКУПЛЯТИ , аю, аеш, <i>рідко</i> ВИКУПАТИ , аю, аеш, <i>недок.</i> , ВИКУПИТИ , плю, пши; <i>мин. ч.</i> ВИКУПЯТЬ ; <i>док.</i>
29.	112.2	6	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121}) + B_{13}(K_{131}) + K_{132}) + C_2(B_{21}(K_{211} + K_{212}))$	ЗЛАЗИТИ , злажу, злазиш, <i>рідко</i> ІЗЛАЗИТИ , лажу, лажш, ЗЛІЗАТИ , <i>рідко</i> ІЗЛІЗАТИ , аю, аеш, <i>недок.</i> , ЗЛІЗТИ , <i>рідко</i> ІЗЛІЗТИ , зу, зеш; <i>мин. ч.</i> зліз і <i>рідко</i> ізліз, ла, ло; <i>док.</i>
30.	111.11	5	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121}) + B_{13}(K_{131})) + C_2(B_{21}(K_{211}) + B_{22}(K_{221}))$	ПРОХОДЖУВАТИСЯ , уюся, уешся, ПРОХОДІТИСЯ , джуся, дшися, <i>розм.</i> ПРОХОДЖАТИСЯ , аюся, аешся, <i>недок.</i> , ПРОЙТИСЯ , пройдуся, пройдешия, <i>розм.</i> ПРОХОДІТИСЯ , ходжуся, ходишия, <i>док.</i>
31.	2.1	3	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112})) + C_2(B_{21}(K_{211}))$	РОЗДУВАТИ , <i>рідше</i> РОЗДИМАТИ , аю, аеш, <i>недок.</i> , РОЗДУТИ , роздую, роздуюш і <i>рідше</i> роздїму, роздїмеш; <i>мин. ч.</i> роздїв, ла, ло; <i>док., що.</i>
32.	2.2	4	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112})) + C_2(B_{21}(K_{211} + K_{212}))$	ДРАПАТИ , ДРЯПАТИ , аю, аеш, <i>недок.</i> , ДРАПНУТИ , ДРЯПНУТИ , ну, неш, <i>док., розм.</i>
33.	2.3	5	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112})) + C_2(B_{21}(K_{211} + K_{212} + K_{213}))$	УМІНАТИ (вмінати), аю, аеш, <i>недок.</i> , УМ'ЯТИ (вм'яти), діал. УМНЯТИ , умнї, умнеш, <i>док., що.</i>
34.	2.4	6	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112})) + C_2(B_{21}(K_{211} + K_{212} + K_{213} + K_{214}))$	УТИНАТИ (втинати), аю, аеш, <i>недок.</i> , УТНУТИ (втнүти), УТЯТИ (втяти), утнї, утнеш, <i>док., кого, що і без дод.</i>

№	Сигна- тура	Кількість слів у ресурсово- му ряді	Структурна формула	Приклад
35.	2.11	4	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + C_2 (B_{21}(K_{211}) + B_{21}(K_{221})))$	НАВЧАТИ, НАУЧАТИ, <i>аю, аєш, недок., НАВЧИТИ, чу, чийш, НАУЧИТИ, учу, учиш, док. кого і без дод.</i>
36.	2.12	5	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + C_2 (B_{21}(K_{211}) + B_{22}(K_{221} + K_{222})))$	ЗГНИВАТИ, <i>рідше</i> ЗОГНИВАТИ, <i>аю, аєш, недок., ЗГНИТИ, згнию, згниєш, рідше ЗГНИТИ / ЗОГНИТИ, що, нещ, док.</i>
37.	2.21	5	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + C_2 (B_{21}(K_{211} + K_{212}) + B_{22}(K_{221})))$	ЗШТОВХУВАТИ, ЗШТОВХУВАТИ, <i>ую, уєш, недок., ЗШТОВХНУТИ, ЗШТОВХНУТИ, ну, неш, розм. ЗШТОВХАТИ, аю, аєш, док., перев.</i>
38.	2.22	6	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + C_2 (B_{21}(K_{211} + K_{222}) + B_{22}(K_{221} + K_{222})))$	ЗАСТРЯВАТИ, ЗАСТРЯГАТИ, <i>аю, аєш, недок., ЗАСТРЯТИ, рідше ЗАСТРЯНУТИ, яну, янеш, ЗАСТРЯТИ, рідше ЗАСТРЯГНУТИ, гну, гнеш; мн. ч. застряв і застряг, ла. ло; док.</i>
39.	2.33	8	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + C_2 (B_{21}(K_{211} + K_{212} + K_{213}) + B_{22}(K_{221} + K_{222} + K_{223})))$	ВПИХАТИСЯ (УПИХАТИСЯ), <i>аюся, аєшся, недок., ВПИХАТИСЯ (ВВПИХАТИСЯ, УВПИХАТИСЯ, УПИХАТИСЯ), аюся, аєшся, док., розм.</i>
40.	2.42	8	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + C_2 (B_{21}(K_{211} + K_{212} + K_{213} + K_{214}) + B_{22}(K_{221} + K_{222})))$	УСИХАТИ (ВСИХАТИ), <i>аю, аєш, недок., УСОХНУТИ (ВСОХНУТИ), УСОХТИ (ВСОХТИ), хну, хнеш, діал. УСХНУТИ (ВСХНУТИ), ну, неш, док.</i>
41.	2.112	6	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + C_2 (B_{21}(K_{211}) + B_{22}(K_{221}) + B_{23}(K_{231} + K_{232})))$	ЗВИВАТИСЯ, <i>рідко</i> ІЗВИВАТИСЯ, <i>аюся, аєшся, недок., ЗВИТИСЯ, зив'юся, зив'єшся, рідко</i> ІЗВИТИСЯ, <i>зів'юся, зів'єшся, діал. ЗВИНУТИСЯ і рідко</i> ЗВИНУТИСЯ, <i>пуюся, нешся, док.</i>
42.	2.1.1	4	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + B_{12}(K_{121})) + C_2 (B_{21}(K_{211}))$	ПРИПРОШУВАТИ, <i>рідко</i> ПРИПРОШУВАТИ, <i>ую, уєш, ПРИПРОШАТИ, аю, аєш, недок., ПРИПРОСИТИ, ошу, осийш, док., кого, перев. з инфін., розм.</i>
43.	2.1.2	5	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + B_{12}(K_{121})) + C_2 (B_{21}(K_{211} + K_{212}))$	ЗДРИГАТИСЯ, <i>рідко</i> ІЗДРИГАТИСЯ, <i>аюся, аєшся, рідше</i> ЗДРИГУВАТИСЯ, <i>уюся, уєшся, недок., ЗДРИГНУТИСЯ, рідко</i> ІЗДРИГНУТИСЯ, <i>пуюся, нешся, док.</i>
44.	2.2.2	6	$\Lambda(X) = C_1 (B_{11}(K_{111} + K_{112}) + B_{12}(K_{121} + K_{122})) + C_2 (B_{21}(K_{211} + K_{212}))$	ВГЛИБЛЮВАТИСЯ (УГЛИБЛЮВАТИСЯ), <i>ююся, юєшся, ВГЛИБЛЯТИСЯ (УГЛИБЛЯТИСЯ), яюся, аєшся, недок., ВГЛИБІТИСЯ (УГЛИБІТИСЯ), вглиб'юся, вглиб'єшся; мн. вглибляється, док., рідко.</i>

№	Сигна-тура	Кількість слів у ресстророваному ряді	Структурна формула	Приклад
45.	22.4	8	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112}) + B_{22}(K_{221} + K_{222})) + C_2(B_{22}(K_{221} + K_{222} + K_{223} + K_{224}))$	ВТЯГУВАТИСЯ (УТЯГУВАТИСЯ) , уюся, уешся, ВТЯГАТИСЯ (УТЯГАТИСЯ) , аюся, аешся, <i>недок.</i> , ВТЯГНУТИСЯ (УТЯГНУТИСЯ) , втягнуся, втягнешся, <i>док.</i> , <i>у що, до чого.</i>
46.	222.2	8	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112}) + B_{12}(K_{121} + K_{122}) + B_{13}(K_{131} + K_{132})) + C_2(B_{21}(K_{211} + K_{212}))$	ВСЛУХАТИСЯ (УСЛУХАТИСЯ) , аюся, аешся, ВСЛУХУВАТИСЯ (УСЛУХУВАТИСЯ) , уюся, уешся, <i>рідко</i> ВСЛУХОВУВАТИСЯ (УСЛУХОВУВАТИСЯ) , уюся, уешся, <i>недок.</i> , ВСЛУХАТИСЯ (УСЛУХАТИСЯ) , аюся, аешся, <i>док.</i> , <i>у що і без дод.</i>
47.	212.111	8	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112}) + B_{12}(K_{121}) + B_{13}(K_{131} + K_{132})) + C_2(B_{21}(K_{211}) + B_{22}(K_{221}) + B_{23}(K_{231}))$	ПРОВОДІТИ, ПРОВАДІТИ , джу, диш, ПРОВОДЖАТИ , аю, аеш, <i>рідко</i> ПРОВОДЖУВАТИ, ПРОВАДЖУВАТИ , ую, уеш, <i>недок.</i> , ПРОВЕСТИ , едү, едеш; <i>мин. ч.</i> провів, велä, лö, <i>рідко</i> ПРОВОДІТИ , воджу, водиш і <i>диал.</i> ПРОВАДІТИ , джу, диш, <i>док.</i>
48.	3.3	6	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121}) + B_{13}(K_{131})) + C_2(B_{21}(K_{211}) + B_{22}(K_{221}) + B_{23}(K_{231}))$	ЗШКРІБАТИ, ЗШКРІБАТИ , аю, аеш, <i>недок.</i> , ЗШКРЕБТИ, ЗШКРЕБТИ, ЗШКРЕБТИ , бү, беш, <i>док.</i> , <i>перех.</i> , <i>розм.</i>
49.	3.11	5	$\Lambda(X) = C_1(B_{11}(K_{111}) + B_{12}(K_{121}) + B_{13}(K_{131})) + C_2(B_{21}(K_{211}) + B_{22}(K_{221}))$	ПЕРЕПЛІСКУВАТИСЯ, ПЕРЕПЛЮСКУВАТИСЯ , уюся, уешся, <i>недок.</i> , ПЕРЕПЛЕСНУТИСЯ , нуся, нешся, ПЕРЕПЛЮСНУТИСЯ , нуся, нешся, <i>док.</i>
50.	4.2	6	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112} + K_{113} + K_{114}) + C_2(B_{21}(K_{211} + K_{212}))$	ВХОДІТИ (УХОДІТИ), ВВІХОДІТИ (УВІХОДІТИ) , джу, диш, <i>недок.</i> , ВВІЙТИ (УВІЙТИ) , увійдү, увійдеш, <i>док.</i>
51.	4.4	8	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112} + K_{113} + K_{114}) + C_2(B_{21}(K_{211} + K_{212} + K_{213} + K_{214}))$	ВТОВКМАЧУВАТИ, ВТОКМАЧУВАТИ (УТОКМАЧУВАТИ) , ую, уеш, <i>недок.</i> , ВТОВКМАЧИТИ (УТОВКМАЧИТИ), ВТОКМАЧИТИ (УТОКМАЧИТИ) , чу, чиш, <i>док.</i> , <i>перех.</i> , <i>розм.</i>
52.	4.22	8	$\Lambda(X) = C_1(B_{11}(K_{111} + K_{112} + K_{113} + K_{114}) + C_2(B_{21}(K_{211} + K_{212}) + B_{22}(K_{221} + K_{222}))$	УДОВОЛЬНЯТИСЯ (ВДОВОЛЬНЯТИСЯ) , заст.

Не здійснюючи докладного лінгвістичного аналізу викладеної моделі та правила «1-2-3-4», зробимо декілька зауважень.

По-перше, модель, очевидно, визначає певну класифікацію на множині українських дієслів. Справді, якщо позначити одержані класи дієслів, марковані поданими в табл. 3.3 сигнатурами, через $\lambda_1, \lambda_2, \dots, \lambda_{52}$, а множини дієслів, що належать до відповідного класу, — через $q(\lambda_1), q(\lambda_2), \dots, q(\lambda_{52})$, то очевидно, що:

$$q(\lambda_i) \cap q(\lambda_j) = \emptyset \text{ при } i \neq j, \quad (3.24)$$

тобто будь-яке з дієслів може належати одному і тільки одному з наведених у табл. 3.3 класів⁴. Це означає, що одержана класифікація є коректною.

Разом з тим, вона є досить точною — показово, що з неї не знайшлося жодного винятку на масиві дієслівних лексем СУМ-11 та СУМ-20.

Отже, можна стверджувати, що одержана класифікація охоплює всі факти з досліджуваного класу *українське дієслово* і репрезентує певний об'єктивний закон для української мови, а не лише для конкретного, хоч і великого Словника.

Природну інтерпретацію введеного закону надає теорія семантичних станів. Припустимо, що існує інтелектуальний механізм, який, аналізуючи будову будь-якої дієслівної лексики X , ідентифікує його семантичний стан $\psi(X)$, тобто встановлює її видові, словозмінні та морфонема-тичні характеристики. Формально дію такого механізму можна зобразити формулою:

$$S \psi(X) = \sigma \psi(X), \quad (3.25)$$

де символом S позначено оператора симетрії, а σ — власне значення цього оператора: $\sigma \in \{(\delta); (\alpha \cdot \beta)\}$. Наприклад, розглянемо дію закону (3.25) на дієслові «ЗСОХНУТИ»:

$$S \psi(\text{ЗСОХНУТИ}) = (\alpha \cdot \beta) \psi(\text{ЗСОХНУТИ}), \quad (3.26)$$

де $\alpha = (1, 2)$; $\beta = (1, 4)$, яка, з одного боку, є наслідком будови відповідної словникової статті:

ЗСИХАТИ, рідко ІЗСИХАТИ, аю, аєш, недок., ЗСОХНУТИ, рідко ІЗСОХНУТИ, ЗСОХТИ, рідко ІЗСОХТИ, хну, хнеш; мин. ч. зсох і зсохнув, ла, ло; док.,

а з іншого — сама визначає цю структуру:

$$(X, Y) \text{ пар } (X, Y) \text{ недок.}, (Z, Q, P, T) \text{ пар } (Z, Q, P, T) \text{ док.} \quad (3.27)$$

⁴ Класифікаційною для дієслів, зрозуміло, є й повна множина із 7140 класів із сигнатурами (α) і $(\alpha \cdot \beta)$.

Текстові відповідники елементів структури в наведених вище термінах виглядають так:

комплекс: $C_1 = \text{ЗСИХАТИ}$, рідко ІЗСИХАТИ , аю, аеш;

блок: $B_{11} = \text{ЗСИХАТИ}$, рідко ІЗСИХАТИ ;

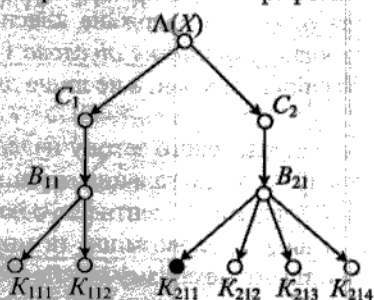
компоненти: $K_{111} = \text{ЗСИХАТИ}$; $K_{112} = \text{ІЗСИХАТИ}$;

комплекс: $C_2 = \text{ЗСОХНУТИ}$, рідко ІЗСОХНУТИ , ЗСОХТИ , рідко ІЗСОХТИ , хну, хнеш; *мин. ч. зсох і зсохнув*, ла, ло;

блок: $B_{21} = \text{ЗСОХНУТИ}$, рідко ІЗСОХНУТИ , ЗСОХТИ , рідко ІЗСОХТИ ;

компоненти: $K_{211} = \text{ЗСОХНУТИ}$; $K_{212} = \text{ІЗСОХНУТИ}$; $K_{213} = \text{ЗСОХТИ}$; $K_{214} = \text{ІЗСОХТИ}$.

Графічний репрезентант структури (α , β) на стані ψ (ЗСОХНУТИ), де $\alpha = (1, 2)$; $\beta = (1, 4)$, зображається таким графом:



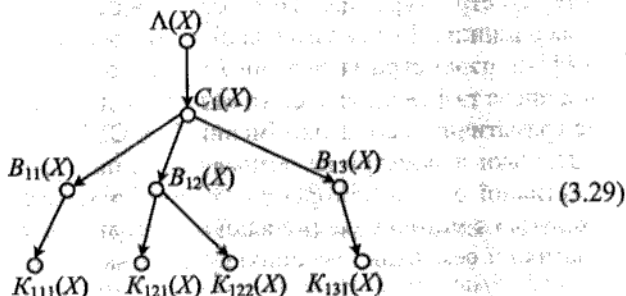
(3.28)

де елемент K_{211} , що відповідає елементу реєстрового ряду ЗСОХНУТИ , затемнено.

Де кілька слів про «порожні» структури класифікації, тобто ті, для яких не знайшлося відповідних дієслів у СУМ-11 та СУМ-20. Згадаймо, наприклад, історію з побудовою Д. І. Менделєєвим класифікації хімічних елементів — число клітин у ній теоретично не обмежене, хоча в часи Д. І. Менделєєва були відомі не більше 63 елементів, та ще й чимало з них були не досить чітко ідентифіковані. У створеній Д. І. Менделєєвим таблиці був шлий ряд пробілів, але система — періодичний закон — дозволяла прогнозувати властивості «пропущених» елементів. Зауважимо, що навіть у наш час відкрито лише 118 елементів, причому останні з них, надважкі, є дуже нестійкими, мають дуже малий час життя, у природі не зустрічаються, а можуть бути одержані лише штучним шляхом на прискорювачах.

Так само і в нашому випадку не виключене існування певних дієслів, які належать до «пропущених» класів. Наявність вільних класів сигналізує про досі не використані словотвірні потенції українського дієслова.

Можливо, більш глибокі дослідження і розвиток мовної системи дозволять виявити такі класи, тим більше, що сама система «прогнозує» їхні морфологічні властивості. Наприклад, серед одноконлексних сигнатур, ідентифікованих у СУМ-11 та СУМ-20, наявна сигнатура (11) — їй відповідає, наприклад, $\Lambda(X) =$ **БУЛЬКОТАТИ**, *очу, очеш, БУЛЬКОТІТИ, очу, отіш, недок.*, а також сигнатура (12) — їй відповідає, наприклад, $\Lambda(X) =$ **РИБАЛИТИ**, *лю, лиш, РИБАЛЧИТИ, рідко РИБАЧИТИ, чу, чиш, недок.* але уже сигнатура (121) — відсутня. З нашої класифікації випливає, що сигнатура (121) визначає видовий комплекс із трьома парадигматичними блоками, у першому з яких є один компонент, у другому — два і в третьому — один компонент. Отже, для елементів цього класу справедливе таке графічне представлення:



і, за аналогією з варіантом сигнатури (12), сигнатурі (121) відповідає така модель $\Lambda(X)$:

$R+\dot{\Lambda}$ ЛИТИ, лю, лиш, $R+\dot{\Lambda}$ ЛЧИТИ і рідко $R+\dot{\Lambda}$ ЧИТИ, чу, чиш, (3.30)
 $R+\dot{\Lambda}$ ЧУВАТИ, ую, уеш, недок.,

де символом R позначено корінь гіпотетичної лексеми.

Таким чином, існує теоретична можливість розвитку лексичної системи у напрямі формування структурних класів, які досі були відсутні в українській мові, але такі, що є дозволені його морфосемантичною системою та узгоджуються із законом прихованої симетрії «1-2-3-4».

З іншого боку, інтуїтивно зрозуміло, що реалізація «надважких» класів, представлених сигнатурами, наприклад (344.444), (444.344), (444.444), дуже малоімовірна — їх просто «не витримає» морфологічна система української мови. Отже, виникає логічне запитання про встановлення меж цієї системи, відповідь на яке можна одержати на шляху проведення комплексних лексико-граматичних і лексико-семантичних досліджень.

3.3. Іменник. Структура та класифікація

У цьому підрозділі застосовано описані вище теоретичні засади теорії Л-системи СУМа як універсального інструмента у виявленні закономірностей морфологічної (граматичної) структури лексем іншої частини мови – іменника.

Факти, встановлені на прикладі українських іменників, свідчать про те, що структура граматичної семантики цього класу слів, як і дієслів, підпорядкована імпліцитним, внутрішньомовним механізмам, певним закономірностям, правилам, які експлікуються у чітко структурованому тексті тлумачного словника [218–221].

Спостереження над корпусом іменникових словникових статей СУМ-11 (всього 52925 іменникових статей) дозволяють зробити висновок про те, що структура $\Lambda(X)$ формується на основі елементарних інформаційних одиниць. Ними є заголовне слово, його лексемні варіанти та лексикографічні параметри заголовного слова (позначення роду, словозмінного класу, числа та функціонального стилю), за допомогою яких здійснюється опис граматичної семантики іменників у СУМ-11 (див. табл. 3.4).

Оскільки параметру на позначення стилю властива наявність певної суб'єктивної оцінки, авторського оцінного характеру, його значення не враховуємо з метою максимальної об'єктивності дослідження структури граматичної семантики іменників української мови.

Таблиця 3.4. Виділення структурних елементів у лівій частині словникової статті СУМ-11

№	Структурний елемент	Зміст структурного елемента	Значення структурного елемента	Реалізація у СУМ-11
1.	X	Заголовне слово	Будь-який іменник, зафіксований у СУМ-11	ВНУК (УНУК і рідко ОНУК)
2.	Парадигма	Фіксування словозмінного класу іменника	1. Параметр однокомпонентний; 2. Параметр однокомпонентний (два варіанти параметра);	бк_а а і у
3.	Рід	Фіксування роду іменника	1. Параметр однокомпонентний; 2. Параметр двокомпонентний; 3. Параметр трикомпонентний;	ч. ж. і с. ч., ж. і с.
4.	Число	Фіксування вживаності іменника у певній числовій формі	1. Параметр однокомпонентний; 2. Параметр двокомпонентний;	збірн., мн. або відсутнє (в одиниці) мн. (одн.) або одн. (мн.)
5.	Стиль	Фіксування функціонально-стильової вживаності іменника	1. Параметр однокомпонентний; 2. Параметр двокомпонентний; 3. Параметр трикомпонентний;	поет. книжн., заст. дорев., розм., зневажл.

Як виявилось в процесі дослідження, функціонування елементарних структуротвірних компонентів $\Lambda(X)$ не є хаотичним і підпорядковується загальним граматичним законам мови, на основі яких творяться типи структур граматичної семантики українських іменників. Закономірне породження таких структур виявляється в ієрархічній організації виявлених структуротвірних компонентів $\Lambda(X)$.

Згідно із засадами теорії Л-системи СУМа, ієрархізацію параметрів граматичної семантики іменників подаємо у вигляді схеми, на якій показано співвідношення формальних граматичних параметрів і умовно названих структурних елементів, розташованих в ієрархічному порядку (див. рис. 3.1).

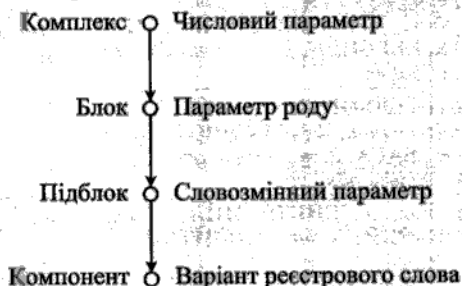


Рис. 3.1. Ієрархізація параметрів граматичної семантики іменників у СУМ-11

Виявлена ієрархічна залежність указує на те, що існує закономірність для всього іменникового корпусу СУМ-11, за якою текст $\Lambda(X)$ для будь-якого X з множини українських іменників однозначно подається у вигляді системи вкладень:

$$\Lambda(X) \supseteq Ч \supseteq Р \supseteq П \supseteq К, \quad (3.31)$$

де символом \supseteq позначено відношення належності, а кожний з чотирьох елементів Ч, Р, П, К являє собою певну частину $\Lambda(X)$, яка має чітко визначену лінгвістичну кваліфікацію: Ч – відображення параметра на позначення числа, Р – відображення параметра на позначення роду, П – відображення параметра на позначення парадигматичного (словозмінного класу), К – відображення параметра на позначення реєстрового слова (та його варіанта).

На найвищому рівні ієрархії, як показано у поданій вище формулі, виділяємо параметр на позначення числа. В ієрархічній залежності кваліфікуємо його як комплекс, а враховуючи параметр, з яким він співвідноситься, називатимемо числовим комплексом. Як показує досліджуваний матеріал СУМ-11, числовий комплекс виділяється за різними значеннями, яких може набувати параметр числа (див. табл. 3.5).

Способи відображення параметра числа у $\Lambda(X)$	Значення, якого набуває параметр числа
<p>ФІЛОСОФ, а, ч.; ФІЛОСОФЕМА, и, ж.; ФЕХТУВАННЯ, я, с.; БЕРІЗКА, <i>рідше</i> БЕРІЗКА, и, ж.; БУСОЛ, БУСЕЛ, сла, ч.; ВОГОНЬ (<i>розм.</i> ОГОНЬ), гніо, ч.; ЗНАХАРКА, <i>рідко</i> ЗНАХОРКА, ЗНАХУРКА, и, ж., <i>заст.</i>; ВНУЧОК (ОНУЧОК, <i>рідко</i> УНУЧОК), чка, ч.</p>	«Звичайне» число (множина не фіксується, але припускається)
<p>БІДНЯЦТВО, а, с., <i>збірн.</i>; ВЗУТТЯ, я, с., <i>збірн.</i>; ВІЛЬШАНИК, у, ч., <i>збірн.</i>, <i>рідко</i>; СМЕРІЧЧЯ, СМЕРЕЧЧЯ, я, с., <i>збірн.</i>; ВЕРЬНИК, ВЕРБНЯК, а, ч., <i>збірн.</i>; БЕСКЕТТА, <i>рідко</i> БЕСКЕДЦЯ, я, с., <i>збірн.</i></p>	Збірне
<p>ГАЛУН¹, а, ч., <i>тільки одн.</i>; ЛЬОН, льону, ч., <i>тільки одн.</i>; ГРАМОТА 1, и, ж., <i>тільки одн.</i>; ПРЕМУДРІСТЬ, <i>рості, ж.</i>, <i>тільки одн.</i>, <i>заст.</i>, <i>уроч.</i>; ВАГА, її, ж., <i>тільки одн.</i>; ТОВКОТНІЧА, і, ж., <i>тільки одн.</i>, <i>розм.</i></p>	Тільки однина
<p>ПЕРИСТОКРИЛКИ, лок, мн.; ВИГЛЯДИНИ, ин, мн., <i>розм.</i>; ВИБОРИ, ів, мн.; ВИВАРКИ, ів, мн.; ПОЖИТКИ, ів, мн., <i>розм.</i></p>	Тільки множина
<p>ПЕТЛЯ 1, ж. (мн. <i>пéтлі, тель</i>); ЧУДО, а, с.; мн. ЧУДА, чуд і ЧУДЕСА, чудес; ТЯМКА, и, ж.; мн. <i>тьмьки, мок; розм.</i>; ЧАГАР, я, ч.; ЧАГАРІ, ів, мн.; ТЯГЛО, а, с.; мн. ТЯГЛА, тягол; ВЕРСТВА, її, ж. (мн. ВЕРСТВИ, <i>верстов</i>); ПОКЛАД¹, у, ч. (мн. ПОКЛАДИ, ів).</p>	Фіксування множини при однині як наслідок ненормативного творення парадигми
<p>РОЗЛІГИ, ів, мн. (<i>одн.</i> розліг, розліг, <i>логу, ч.</i>); КЕНДІ, КИНДІ, ів, мн. (<i>одн.</i> КЕНДЯ, КИНДЯ, і, ж.), <i>діал.</i>; ТРОПІЗМИ, ів, мн. (<i>одн.</i> ТРОПІЗМ, у, ч.); КРУЧЕНИКИ, ів, мн. (<i>одн.</i> КРУЧЕНИК, а, ч.); КУЧЕРІ, ів, мн. (<i>одн.</i> КУЧЕР, я, ч.); КУЧЕРИКИ, ів, мн. (<i>одн.</i> КУЧЕРИК, а, ч.); РУНА³, рун, мн. (<i>одн.</i> РУНО, а, с.).</p>	Фіксування форм(и) однини при множині: а) як частіше вживаної форми
<p>ПАХОЩІ, ів, мн. (<i>рідко одн.</i> ПАХОШ, у, ч. і ж.); БЛИЗНЯТА, ят, мн. (<i>рідко одн.</i> БЛИЗНЯ, яти, с.); ЧАТИ, чаг, мн., <i>рідше одн.</i> ЧАТА, и, ж.; ВЕРЕДИ, ів, мн. (<i>рідко одн.</i> ВЕРЕД, у, ч.); ДІТЛАХИ, ів, мн. (<i>рідко одн.</i> ДІТЛАХ, а, ч.), <i>розм.</i> ДІТОНЬКИ, ньок, мн. (<i>рідко одн.</i> ДІТОНЬКА, и, ж.); СВЯТОЩІ, ів, мн. (<i>рідко одн.</i> СВЯТОЩ, і, ж.); ШАТИ, шат, мн. (<i>рідко одн.</i> ШАТА, и, ж.), <i>поет.</i>, <i>рит.</i></p>	б) як раритетних, рідковживаних форм із стильовим маркуванням
<p>ПОЛІЩУКИ, ів, мн. (<i>одн.</i> ПОЛІЩУК, а, ч.; ПОЛІЩУЧКА, и, ж.); НОРВЕЖЦІ, ів, мн. (<i>одн.</i> НОРВЕЖЕЦЬ, жця, ч.; НОРВЕЖКА, и, ж.); МУЛАТИ, ів, мн. (<i>одн.</i> МУЛАТ, а, ч.; МУЛАТКА, и, ж.); УДІ, мн. (<i>одн.</i> УДІ, ч. і ж.), <i>невідм.</i>; ПЕЙЗАНИ, ів, мн. (<i>одн.</i> ПЕЙЗАН, ПЕЙЗАНИН, а, ч.); ПЕЙЗАНКА, и, ж.); АБХАЗЦІ, ів, мн. (<i>одн.</i> АБХАЗ, а і АБХАЗЕЦЬ, зця, ч.; АБХАЗКА, и, ж.).</p>	Фіксування форм однини і множини у $\Lambda(X)$ іменників – на позначення народностей

Тексти різних комплексів у одній словниковій статті не перетинаються — вони є чітко розмежованими, отже, для елемента \mathcal{C} справедливим є формальне представлення:

$$\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2,$$

де символами \mathcal{C}_1 та \mathcal{C}_2 позначено перший і другий (якщо він є в аналізованій словниковій статті) числові комплекси, причому $\mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset$, тобто їхній перетин порожній. Наприклад, якщо:

$$\Lambda(X) = \text{ТЕРМИ, терм, мн.},$$

то елемент \mathcal{C} містить лише один числовий комплекс \mathcal{C}_1 :

$$\mathcal{C}_1 = \mathcal{C} = \Lambda(X) = \text{ТЕРМИ, терм, мн.}$$

Виділення числових комплексів здійснюється відповідно до відображення у словниковій статті співвідносного з комплексом параметра на позначення числа. Розглянемо приклади виділення числових комплексів у межах словникової статті $\Lambda(X)$ іменників української мови.

Словникові статті, що визначаються одним числовим комплексом, умовно відображаємо як $\Lambda(X) = \mathcal{C}_1$, де \mathcal{C}_1 може набувати різних числових значень: **БРАТКИ**, ів, мн., де $\mathcal{C}_1 = \text{БРАТКИ, ів, мн.}$; **АПАРАТУРА**, и, жс., збірн., де $\mathcal{C}_1 = \text{АПАРАТУРА, и, жс., збірн.}$; **ПРЕМУДРІСТЬ**, рості, жс. тільки одн., заст., уроч., де $\mathcal{C}_1 = \text{ПРЕМУДРІСТЬ, рості, жс. тільки одн., заст., уроч.}$

Як свідчать наведені приклади, числовий комплекс реалізується з урахуванням різних значень числового параметра, відображеного за допомогою відповідної системи ремаркування, представленої в СУМ-11.

Словникові статті, що складаються з двох числових комплексів, умовно позначаємо як $\Lambda(X) = \mathcal{C}_1 + \mathcal{C}_2$:

ІНКУНАБУЛИ, ул, мн. (одн. **ІНКУНАБУЛА**, и, жс.), де $\mathcal{C}_1 = \text{ІНКУНАБУЛИ, ул, мн.}$; $\mathcal{C}_2 = \text{одн. ІНКУНАБУЛА, и, жс.}$;

АЕРОБИ, ів, мн. (одн. **АЕРОБ**, а, ч.), біол., де $\mathcal{C}_1 = \text{АЕРОБИ, ів, мн., біол.}$; $\mathcal{C}_2 = \text{одн. АЕРОБ, а, ч., біол.}$;

ТЮФТЕЛЬКИ, льок, мн. (одн. **ТЮФТЕЛЬКА**, и, жс.), де $\mathcal{C}_1 = \text{ТЮФТЕЛЬКИ, льок, мн.}$; $\mathcal{C}_2 = \text{одн., ТЮФТЕЛЬКА, и, жс.}$;

АЛЬВЕОЛИ, бл, мн. (одн. **АЛЬВЕОЛА**, и, жс.), анат., де $\mathcal{C}_1 = \text{АЛЬВЕОЛИ, бл, мн., анат.}$; $\mathcal{C}_2 = \text{одн. АЛЬВЕОЛА, и, жс., анат.}$;

УКРАЇНЦІ, ів, мн. (одн. **УКРАЇНЕЦЬ**, нця, ч.; **УКРАЇНКА**, и, жс.), де $\mathcal{C}_1 = \text{УКРАЇНЦІ, ів, мн.}$; $\mathcal{C}_2 = \text{одн. УКРАЇНЕЦЬ, нця, ч.; УКРАЇНКА, и, жс.}$;

ЧУДО, а, с.; *мн.* **ЧУДА**, чуд / **ЧУДЕСА**, чудес, де $Ч_1 = \text{ЧУДО}$, а, с.; $Ч_2 = \text{мн. ЧУДА}$, чуд / **ЧУДЕСА**, чудес;

ШАТИ, шат, *мн.* (рідко одн. **ШАТА**, и, ж.), *поет., рит.*, де $Ч_1 = \text{ШАТИ}$, шат, *мн.*, *поет., рит.*; $Ч_2 = \text{рідко одн. ШАТА}$, и, ж., *поет., рит.*

Важливою передумовою вивчення структури граматичної семантики іменників у їх відображенні СУМ-11 є повне відмежування від будь-яких лексикографічно зумовлених лінгвістичних фактів і виявлення об'єктивних закономірностей у системі мови. Завдання тлумачного словника полягає у фіксуванні та висвітленні саме таких мовних фактів, однак при цьому завжди простежуються вироблені практикою словникарства лексикографічно зумовлені лінгвістичні факти, які при виведенні закономірностей із тексту словникової статті необхідно враховувати і обумовлювати. Тому для встановлення параметризації граматичної семантики іменників у СУМ-11 першочергове завдання полягає в тому, щоб виявити й відповідно кваліфікувати лінгвістичну інформацію, зумовлену способами лексикографування.

Як показує аналізований матеріал, відображення числового параметра у словникових статтях зумовлене лексикографічними правилами укладання СУМ-11. Згідно з вимогами, зафіксованими у передмові СУМ-11, відображення числа обумовлюється такими лексикографічними правилами [85]:

1. Відмінювані слова вносяться до реєстру Словника в початковій граматичній формі: іменники, кількісні числівники та займенники не-прикметникової форми – в називному відмінку однини. Напр.: Двір, двора і двору, ч.; Річка, и, ж.; Море, я, с. При аналізі такого способу лексикографування граматичної семантики іменників у СУМ-11 стає очевидною невідповідність числового ремаркування структурі граматичної семантики цих іменників: наявність парадигми в множині не відображена у словниковій статті, а лише припускається. Відтак виникає необхідність урахування зазначеного способу лексикографування і наступної кваліфікації такої структури граматичної семантики іменників як двочислової, з корелятивними формами однини і множини: $\Lambda(X) = Ч_1(\text{одн.}) + Ч_2(\text{мн.})$.

2. Іменники, частіше вживані в множині, подаються в цій формі, а в дужках наводиться форма однини з відповідним граматичним оформленням: БЛИЗНЯТА, ят, *мн.* (рідко одн. БЛИЗНЯ, яти, с.). За таким зразком оформлюються й назви народів.

Прикладом такого способу лексикографування є іменники: ШАТИ, шат, *мн.* (рідко одн. ШАТА, и, ж.), *поет., рит.*; ПАХОЩІ, ів, *мн.* (рідко одн. ПАХОЩ, у, ч. і ж.); КЕНДІ, КИНДІ, ів, *мн.* (одн. КЕНДЯ, КИНДЯ, і, ж.), *діал.*

ШКАРБАНІ, ШКАРБУНІ, ів, мн. (одн. **ШКАРБАН, ШКАРБУН**, а, ч.), розм.;
ВУСА, ВУСИ, ів, мн. (одн. **ВУС**, рідко **УС**, а, ч.).

Проаналізувавши ліві частини словникових статей, розроблені укладачами згідно з цим правилом, приходимо до висновку про їх двокомплексну структуру: $\Lambda(X) = \mathcal{C}_1(\text{мн.}) + \mathcal{C}_2(\text{одн.})$, що визначається наявністю одночасно двох ремарок числа: *мн.* і *одн.* Особлива структура граматичної семантики цього класу слів з послідовним фіксуванням множини й одиници зумовлена лише функціонально-стильовим параметром, унаслідок чого такі іменники вносяться до реєстру у формі множини. Враховуючи об'єктивну структуру граматичної семантики описаного типу слів приходимо до висновку, що такі словникові статті є прикладом двочислової структури з наявними граматично не маркованими парадигмами одиниці і множини. Тому такі іменники відносимо до класу лексем зі звичайною, граматично немаркованою двочисловою структурою з корелятивними формами одиниці і множини: $\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$. Класи іменників, зумовлені правилами 3 та 4, також визначаються подібною структурою граматичної семантики.

3. Іменники, які нормативно вживаються і в одиниці, і в множині, але форма множини стилістично маркована (за допомогою стильових ремарок *рідко, перев., рідше*), фіксуються в одиниці, а в дужках наводиться форма множини: **ПОХОРОН**, у, ч. (*рідше мн. ПОХОРОНИ*, ів); **ЧОРНОГОЛОВЕЦЬ**, вця, ч. (*перев. мн. ЧОРНОГОЛОВЦІ*, ів), *заст.*; **ЧОРНОБРИВЕЦЬ**, вця, ч. (*перев. мн. ЧОРНОБРИВЦІ*, ів), *діал.*; **ТОВЧЕНІК**, а, ч., *перев. мн. ТОВЧЕНІКИ*, ів.

4. Іменники, які мають форми одиниці і множини, але в множині як стилістично марковані характеризуються ненормативною відмінковою парадигмою (іменники, у лівих частинах словникових статей яких фіксування форми множини у дужках зумовлене як стилістичною маркованістю, так і ненормативною відмінковою парадигмою): **ПОРІЧКА**, и, ж. (*перев. мн. ПОРІЧКИ*, чок); **СУНИЦЯ**, і, ж. (*перев. мн. СУНИЦІ, ніць*); **СУНІЧКА**, и, ж. (*перев. мн. СУНИЧКИ*, чок); **ПОРІЧЕЧКА**, и, ж. (*перев. мн. ПОРІЧЕЧКИ*, чок);

або лексемними варіантами вихідної форми у множині і ненормативною відмінковою парадигмою: **ОКО** а, с. (*мн. ОЧІ, рідко ВІЧІ, очей* – *перев. з прийм. у*); **ГРА**, и, ж. (*мн. ІГРИ, ігор*);

або варіантами вихідної форми у множині, зумовленими акцентуаційними особливостями і ненормативною відмінковою парадигмою: **ВЕРСТВА**, и, ж. (*мн. ВЕРСТВИ, вёрств*); **ПЕТЛЯ**, і, ж. (*мн. ПЕТЛІ, тель*); **РОЗКІШ**, кош, ж. (*мн. РОЗКОШІ*, ів); **ПЕРО**, а, с. (*мн. ПЕРА*, пер).

Подібні структури лівої частини словникових статей іменників розглядатимемо, враховуючи не спосіб їх представлення у СУМ-11, а керуючись об'єктивною структурою їх граматичної семантики, тобто як такі, що мають дві співвідносні числові граматично немарковані парадигми в однині і множині та визначаються формулою:

$$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.}).$$

Таким чином, виявлені лексикографічно зумовлені типи структур граматичної семантики іменників української мови було умовно трансформовано для подальшого автоматичного аналізу лівих частин словникових статей досліджуваних лексем (див. табл. 3.6).

Таблиця 3.6. Трансформація лексикографічно зумовлених структур $\Lambda(X)$ у СУМ-11

Відображення лексикографічно зумовлених структур $\Lambda(X)$	Формула, якою визначається трансформований структурний тип
КЕНЦІ, КИНЦІ, ів, мн. (одн. КЕНДЯ, КИНДЯ, і, жс.), діал.;	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$
ВРУНА (УРУНА), врун, мн. (одн. ВРУНО, а, е.), с. г.;	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$
ПАХОЩІ, ів, мн. (рідко одн. ПАХОЩ, у, ч. і жс.);	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$
БЛИЗНЯТА, ят, мн. (рідко одн. БЛИЗНЯ, яти, с.);	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$
ОРЕЛІ, орель, мн. (одн. ОРЕЛЬ і ОРЕЛЯ, і, жс.), заст.;	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$
КУСТРИ, ів, мн. (одн. КУШТРА і КУШТРА, і, жс.), діал.;	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$
ІНЦІАЛИ, ів, мн. (одн. ІНЦІАЛ, а, ч.);	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$
ОКО ¹ а, с. (мн. ОЧІ, рідко ВІЧІ, очей);	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$
ХАЗЯЇН, а, ч. (мн. ХАЗЯЇНИ, ів і ХАЗЯЇ, ів);	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$
ПЕРО, а, с. (мн. ПЕРА, пер).	$\Lambda(X) = \mathcal{C}_1(\text{одн.}) + \mathcal{C}_2(\text{мн.})$

У результаті вивчення структури лівої частини словникової статті іменників на позначення народів, що, згідно з правилами лексикографування СУМ-11, подаються у множині як вихідній формі з паралельною фіксацією в дужках форм однини, можемо зробити висновок, що така структура словникової статті пов'язана не з лексикографічними вимогами, а відображає семантичну спорідненість об'єднаних нею лексем, вказує на зв'язки між формами однини і множини даних іменників, на співвідносність лексем, вжитих у множині та семантично вивідних з них іменників у однині. Тому виділення такої структури словникової статті як двочислової вважаємо лінгвістично виправданим фактом.

Граматична категорія числа іменника, реалізуючись системою відмінкових форм, окремих для однини і множини, пов'язана з іншими іменниковими категоріями – роду та відмінка, взаємодія з якими передається за допомогою спеціальних морфологічних показників. Це дає підставу для кваліфікації числового параметра як структурного елемента найвищого рівня, у межах якого і від значень якого змінюються набір та якісна характеристика словозмінних показників.

Наступним структурним рівнем ієрархізації є блок, з яким співвідноситься параметр на позначення роду. Друга позиція параметра роду у виведеній ієрархії структурних елементів лівої частини іменникової статті зумовлена тим, що родова належність, відображена в системі морфологічних показників, є вищою щодо наступних структурних елементів, з якими співвідносяться парадигматичний (словозмінний) параметр та лексемні варіанти заголовного слова, але нижчою щодо числового комплексу. Визнання показника параметра роду як структурного елемента вищого рангу щодо показника парадигматичного класу ґрунтується на тому, що значення парадигматичного класу формується на основі поняття роду й визначається ним (див. табл. 3.7).

Таблиця 3.7. Вплив показника роду на словозмінний параметр іменника

Словникова стаття $\Lambda(X)$	Вплив показника роду на словозмінний параметр іменника
ГАМАЗЕЙ, ю, ч., ГАМАЗЕЯ, і, ж., заст.;	ГАМАЗЕЙ (ч.) → словозм. параметр (-ю); ГАМАЗЕЯ (ж.) → словозм. параметр (-ї);
ДРОЖ, розм. ДРЕЖ, дрѡжу, ч. і дрѡжі, ж.;	ДРОЖ (ч.) → словозм. параметр (дрѡжу); ДРОЖ (ж.) → словозм. параметр (дрѡжі); ДРЕЖ (ч.) → словозм. параметр (дрѡжу); ДРЕЖ (ж.) → словозм. параметр (дрѡжі);
ОВОЧ, заст. ОВОЩ, у, ч. і рідко і, ж.;	ОВОЧ (ч.) → словозм. параметр (-у); ОВОЧ (ж.) → словозм. параметр (-ї); ОВОЩ (ч.) → словозм. параметр (-у); ОВОЩ (ж.) → словозм. параметр (-ї);
НОТАБЕНА, в, ж., рідко НОТАБЕНЕ, с., невідм.;	НОТАБЕНА (ж.) → словозм. параметр (-и); НОТАБЕНЕ (с.) → словозм. параметр (невідм.);
ЖИВОКІСТ, живокѡсту, ч., розм. ЖИВОКОСТЬ, кості, ж. (Symphutum L.);	ЖИВОКІСТ (ч.) → словозм. параметр (живокѡсту); ЖИВОКОСТЬ (ж.) → словозм. параметр (-кості);
БОРТЬ, і, ж. і я, ч., заст.;	БОРТЬ (ж.) → словозм. параметр (-ї); БОРТЬ → словозм. параметр (-я);
ПРОДАЖ, у, ч. і рідко і, ж.;	ПРОДАЖ (ч.) → словозм. параметр (-у); ПРОДАЖ (ж.) → словозм. параметр (-ї);
ЛАЗУР, і, ж., рідко у, ч.	ЛАЗУР (ж.) → словозм. параметр (-ї); ЛАЗУР (ч.) → словозм. параметр (-у);

Параметр роду позначається за допомогою лексикографічної ре-марки, співвідносно з родом, яка обов'язково набуває певних кількісних характеристик. Залежно від цього виділяємо такі типи словникових статей:

1. Іменникові статті з одним параметром роду: *ч.*, або *ж.*, або *с.* Виділення цього структурного елемента зумовлюється межами числового комплексу, що веде до формального відображення:

$$\Lambda(X) = (Ч_1(P_1)),$$

де дужки () вказують на входження параметра роду P_1 до складу кожного числового комплексу $Ч_1$, виявленого в даній іменниковій статті $\Lambda(X)$.

Наприклад: БОТАНІК, *а. ч.*; ГАРКЕБУЗ, АРКЕБУЗ, *а. ч.*; БЕСКІД, БЕСКЕД, БЕСКЕТ, *у. ч.*; БОРОТЬБА, *и. ж.*; БАКЛАГА, БОКЛАГА, *и. ж.*; ЗНАХАРКА, *рідко* ЗНАХОРКА, ЗНАХУРКА, *и. ж., заст.*; ГАРМОНІЧ-НІСТЬ, *ності, ж., рідко*; БОТВІННЯ, *я. с., збірн.*; БАЗАРУВАННЯ, *рідше* БАЗАРІОВАННЯ, *я. с., розм.*; ВНУЧАТКО (ОНУЧАТКО, *рідко* УНУ-ЧАТКО), *а. с.*

2. Іменникові статті з двома параметрами роду: *ч. і ж.*, *ч. і с.*, *ж. і ч.*, *ж. і с.*, *с. і ч.*, *с. і ж.*, *ч. і рідко с.*, та *ч. і рідко с.*, що, відповідно до системи вкладення в межах числового комплексу $Ч_1$, визначаються формулою:

$$\Lambda(X) = (Ч_1(P_1+P_2)),$$

де P_1 та P_2 – параметри роду, виділені в межах кожного числового комплексу $Ч_1$ іменниковій статті $\Lambda(X)$. Наприклад: ХИТРІЮГА, *и. ч. і ж., розм.*; ХЛОПЧИСЬКО, *а. ч. і с., розм.*; БІДНЯЖЕЧКА, *и. ж. і ч.*; КОНТРАЛЬТО, *невідм., с. і ч.*; БІДОЛАХА, *рідко* БІДОЛАГА, БІДОЛАКА, *и. ч. і ж., розм.*; СЕРДЕГА, *рідко* СЕРДЕКА, СЕРДЯГА, *и. ч. і рідше ж.*

3. Іменникові статті з трьома параметрами роду: *ч.*, *ж. і с.*, що, відпо-відно до системи вкладення у межах числового комплексу, визначаються формулою:

$$\Lambda(X) = (Ч_1(P_1 + P_2 + P_3)),$$

де параметри роду P_1 , P_2 та P_3 виділяються в межах кожного числового комплексу $Ч_1$ іменниковій статті $\Lambda(X)$. Наприклад: ЛАДО, *а. ч., ж. і с., нар.-поет.*; ЗАБУДЬКО, *а. ч., с. і ж., розм.*

Виділений на основі параметра роду блок як структурний елемент іменниковій статті $\Lambda(X)$ пов'язаний зі словозмінним параметром, що формує структурний елемент нижчого рівня – підблок.

Ми зафіксували словникові статті з різним кількісним розподілом цього параметра. Керуючись виведеним правилом – єдиним для всього імен-

никового корпусу СУМ-11, за яким текст $\Lambda(X)$ для будь-якого X з множини українських іменників однозначно подається у вигляді системи вкладень (див. рис. 3.1), виділення парадигматичних (словозмінних) підблоків здійснюватиметься за відповідними формулами:

1. Іменникові статті з одним словозмінним параметром у межах блоку визначаються формулою:

$$\Lambda(X) = (Ч_1(P_1(\Pi_1))),$$

де парадигматичний підблок Π_1 виділяється в межах блоку P_1 , що входить до числового комплексу $Ч_1$ будь-якої іменникової статті $\Lambda(X)$. Наприклад: **БЛАГОДІЙНІСТЬ**, ності, *ж.*; *dorev.*; **ЗАМІРЮВАННЯ**, я, *с.*, *спец.*; **ГЕНЕЗИС**, у, *ч.*; **САМІТНИК**, *рідше САМОТНИК*, а, *ч.*; **УРОДЖЕ-НІСТЬ (ВРОДЖЕНІСТЬ)**, ності, *ж.*; **ДЕХКАН**, **ДЕХКАНИН**, а, *ч.*; **ВНУЧЕНЬКА (ОНУЧЕНЬКА, рідко УНУЧЕНЬКА)**, и, *ж.*; **ПОРОМ**, *рідко ПАРОМ*, *заст.* **ПОРОН**, а, *ч.*; **ЗНАХАРСТВО**, *рідко ЗНАХОРСТВО, ЗНАХУРСТВО*, а, *с.*, *заст.*

2. Іменникові статті з двома словозмінними параметрами в межах блоку визначаються формулою:

$$\Lambda(X) = (Ч_1(P_1(\Pi_1 + \Pi_2))),$$

де парадигматичні підблоки Π_1 та Π_2 виділяються в межах блоку P_1 , що входить до числового комплексу $Ч_1$ будь-якої іменникової статті $\Lambda(X)$. Наприклад: **ЗНАХАР**, я, *рідко ЗНАХОР, ЗНАХУР*, а, *ч.*, *заст.*; **ЯТЕЛ**, **ЯТІЛ**, ятла, **ЯТІЛЬ**, ятеля, *ч.*, *діал.*; **САЛГАН**, а, **САЛОГАН**, **САЛОГОН**, а, *ч.*, *діал.*; **ДЕБРА**, и, *рідко ДЕБР, ДЕБРЬ*, і, *ж.*, *зах.*; **ГЛЯНЕЦЬ**, *нр.*, **ГЛЯНС**, **ГЛЯНЦ**, у, *ч.*

3. Іменникові статті з трьома словозмінними параметрами в межах блоку визначаються формулою:

$$\Lambda(X) = (Ч_1(P_1(\Pi_1 + \Pi_2 + \Pi_3))),$$

де парадигматичні підблоки Π_1 , Π_2 та Π_3 виділяються в межах блоку P_1 , що входить до числового комплексу $Ч_1$ будь-якої іменникової статті $\Lambda(X)$. Наприклад: **КАДІБ**, доба, **КАДІВБ**, довба, *рідко КАДУБ*, а, *ч.* (єдина лексема з такою структурою).

Згідно з теорією лексикографічних систем, меншим за підблок структурним елементом виступає компонент лівій частини іменникової статті. Таким компонентом є ресстрове слово та його морфологічні, фонетичні чи акцентуаційні варіанти, які об'єднуються спільним лексичним значенням, але відмінні за певними граматичними параметрами. Виділення компо-

нента словникової статті також відбувається відповідно до виведеної ієрархічної вкладеності структурних елементів (див. рис. 3.1).

У лівих частинах іменникових статей встановлено такий кількісний розподіл компонентів:

1. Іменникові статті з одним компонентом, що визначається формулою:

$$\Lambda(X) = (Ч_1(P_1(\Pi_1(K_1))))),$$

де парадигматичний підблок Π_1 виділяється в межах блоку P_1 , що входить до числового комплексу $Ч_1$ із реєстровим словом K_1 будь-якої іменникової статті $\Lambda(X)$. Наприклад: СМІШЕЧКИ, чок, *мн.*; ДОБРІДІЙ, я, *ч.*, *заст.*; ХМИЗЯ, я, *с.*, *збірн.*; ЧЕРЕШЕНЬКА, и, *жс.*; ЩЕДРУВАННЯ, я, *с.*; ЩЕБЕТУНЧИК, а, *ч.*

2. Іменникові статті з двома компонентами, що визначається формулою:

$$\Lambda(X) = (Ч_1(P_1(\Pi_1(K_1 + K_2))))),$$

де парадигматичний підблок Π_1 виділяється в межах блоку P_1 , що входить до числового комплексу $Ч_1$ із реєстровим словом K_1 та його лексемним варіантом K_2 будь-якої іменникової статті $\Lambda(X)$. Наприклад: ПОЛЯДВИЦЯ, ПОЛЕНДВИЦЯ, і, *жс.*; ЗАКАМ'ЯНІЛІСТЬ, *рідше* ЗАКАМЕНІЛІСТЬ, *лесті*, *жс.*; ВСИПИЩЕ (УСИПИЩЕ), а, *с.*; ГОМУНКУЛ, ГОМУНКУЛУС, а, *ч.*; БЕРЕЗКА¹ *рідше* БЕРІЗКА, и, *жс.*; УРЯД, *заст.* УРЯД¹ у, *ч.*; СХРОПУВАННЯ, *рідше* СХРАПУВАННЯ, я, *с.*; ПІДІГРІВ, *рідко* ПЦОГРІВ, у, *ч.*

3. Іменникові статті з трьома компонентами, що визначається формулою:

$$\Lambda(X) = (Ч_1(P_1(\Pi_1(K_1 + K_2 + K_3))))),$$

де парадигматичний підблок Π_1 виділяється в межах блоку P_1 , що входить до числового комплексу $Ч_1$ із реєстровим словом K_1 та його лексемними варіантами K_2 та K_3 будь-якої іменникової статті $\Lambda(X)$. Наприклад: ВНУКА (ОНУКА, *рідко* УНУКА), и, *жс.*; СЕМИРЯЖКА, СЕМЕРЯЖКА, СЕМРЯЖКА, и, *жс.*, *заст.*; ДЗБАН, ДЖБАН, *розм.* ЖБАН, а, *ч.*; СЕСТРІНИЦЯ, СЕСТРИНИЦЯ, СЕСТРІНИЦЯ, і, *жс.*, *діал.*; ЗНАХАРКА, *рідко* ЗНАХОРКА, ЗНАХУРКА, и, *жс.*, *заст.*; БЕСКИД, БЕСКЕД, БЕСКЕТ, у, *ч.*; ВИШКРЕБКИ, ВІСКРЕБКИ, ВІСКРІБКИ, ів, *мн.*, *розм.*

Проведений аналіз граматичної семантики іменників у словникової репрезентації дозволив за аналогією із дієсловом вивести комплексні правила побудови структури граматичної семантики іменників української мови, найважливішими з яких є такі:

1. Кожний іменник у СУМ-11 представлений словом з фіксованою лексичною семантикою за наявності одного або двох числових значень.

Наприклад, у словнику слово **СТЕКЛЯРУС**, *у, ч.* збірн. із значенням «Тоненькі короткі трубочки з кольорового скла, що їх використовують для намиста, оздоблення тканини, одягу і т. ін.; різновид бісеру» визначається однією числовою формою – однини, якою формально реалізується категорія збірності.

Слово **КУЧЕРЯВЧИКИ**, *ів, мн.*, *рідко* із значенням «кучеряве або завите волосся» реалізується й визначається у словнику однією числовою формою – множини.

Одночасно дві числові форми – однини та множини відображені в словникових статтях при лексемах: **ДІТЛАХИ**, *ів, мн.* (*рідко одн.* **ДІТЛАХ**, *а, ч.*), *розм.*; **ВОДОХРЕЩЕ**, *а, с.*; **ВОДОХРЕЩІ**, *ів, мн.*; **ПАНЧОХИ**, *чіх, мн.* (*одн.* **ПАНЧОХА**, *и, ж.*); **ПАЛЬЧАТА**, *чат, мн.* (*одн.* **ПАЛЬЧА**, *ати, с.*), а також припускаються у таких словникових статтях: **РАЦІОНАЛІЗМ**, *у, ч.*; **ПАНІСТКА**, *и, ж.*; **АЛЬПАГА**, **АЛЬПАКА**, *і, ж.*; **ЗЕРНЯТКО**, *розм.* **ЗЕРНЯТКО**, *а, с.*; **ВНУЧКА** (**ОНУЧКА**, *рідко* **УНУЧКА**), *и, ж.*

У словниковій статті **ОДНОСЕЛЬЧАНИ**, **ОДНОСІЛЬЧАНИ**, *чан, мн.* (*одн.* **ОДНОСЕЛЬЧАНИН**, **ОДНОСІЛЬЧАНИН**, *а, ч.*; **ОДНОСЕЛЬЧАНКА**, **ОДНОСІЛЬЧАНКА**, *и, ж.*) також наявні два числових комплекси: \mathcal{C}_1 : **ОДНОСЕЛЬЧАНИ**, **ОДНОСІЛЬЧАНИ**, *чан, мн.*; \mathcal{C}_2 : *одн.* **ОДНОСЕЛЬЧАНИН**, **ОДНОСІЛЬЧАНИН**, *а, ч.*; **ОДНОСЕЛЬЧАНКА**, **ОДНОСІЛЬЧАНКА**, *и, ж.* При цьому елементи \mathcal{C}_1 та \mathcal{C}_2 об'єднують наявні менші структурні компоненти статті – блок, підблок та компонент ремарками *одн.* чи *мн.*

Це правило регламентує функціонування числового параметра у структурі граматичної семантики іменників української мови.

2. Кожний іменник у СУМ-11, представлений словом з фіксованою лексичною семантикою та конкретним значенням числа, може характеризуватися не більше ніж трьома різними значеннями роду. Виняток становлять множинні іменникові лексеми, значення роду в яких відсутні.

Наприклад, іменникова стаття **ВИЗВІЛ**, *візволу, рідше визво́лу, ч.* формується одним рід-блоком, що набуває значення чоловічого роду.

У лівій частині іменникової статті **БІДА́ХА**, *рідко* **БІДА́ГА**, *и, ч.* і *ж.* реалізуються два рід-блоки (спарені): \mathcal{P}_1 : **БІДА́ХА**, *рідко* **БІДА́ГА**, *и, ч.*;

\mathcal{P}_2 : **БІДА́ХА**, *рідко* **БІДА́ГА**, *и, ж.*

У лівій частині іменникової статті **ЗАБУДЬКО**, а, ч., с. і ж., розм. реалізуються три рід-блоки: P_1 : **ЗАБУДЬКО**, а, ч., розм.; P_2 : **ЗАБУДЬКО**, а, с., розм.; P_3 : **ЗАБУДЬКО**, а, ж., розм.

При цьому елементи в межах P_1 об'єднуються в рід-блок ремаркою ч., елементи P_2 – ремаркою с., а елементи P_3 – ремаркою ж.

3. Кожний іменник, представлений у СУМ-11 словом з фіксованою лексичною семантикою, конкретним значенням числа і вказівкою на родову ознаку, може належати не більше ніж до трьох парадигматичних (словозмінних) класів.

Наприклад, у лівій частині словникової статті **ВІДЛІТ**, льоту і лѣту, ч. виділяється один парадигматичний підблок, що характеризується двома варіантами парадигматичного значення.

У статті **КУПІЛЬ**, купелі, ж. і купело, ч. виділяємо по одному парадигматичному підблоку у кожному з двох рід-блоків: P_1 : **КУПІЛЬ**, купелі, ж. та P_2 : **КУПІЛЬ**, купело, ч.

У іменниковій статті **САЛГАН**, а, **САЛОГАН**, **САЛОГОН**, а, ч., діал. у межах одного числового комплексу і одного блоку реалізуються два парадигматичних підблоки: 1) **САЛГАН**, а, ч. і 2) **САЛОГАН**, **САЛОГОН**, а, ч.

З трьома парадигматичними параметрами в СУМ-11 зафіксовано унікальну словникову статтю: **КАДІВЬ**, доба, **КАДІВЬ**, довба, рідко **КАДУБ**, а, ч., де в межах одного числового комплексу та одного рід-блоку реалізуються три парадигматичних підблоки.

4. Кожний іменник, представлений у СУМ-11 словом з фіксованою лексичною семантикою, конкретним значенням числа, вказівкою на родову ознаку і співвіднесеністю з конкретним словозмінним класом, може мати не більше трьох лексемних варіантів.

Утворення варіантних компонентів реєстрової лексеми пов'язане з морфологічними (префіксальними, суфіксальними, префіксально-суфіксальними, кореневими та коренево-суфіксальними), фонетичними чи акцентуаційними змінами реєстрового слова.

Наприклад, словникові статті **ТОВАРИСТВО**, а, с.; **СОЛЬФЕДЖІО**, невід., с., муз.; **УСПІХ**, у, ч.; **ВЕЛИЧ**, і, ж. та ін. містять лише один компонент, який стосується одного числового комплексу, одного рід-блоку й одного парадигматичного підблоку.

У іменникових статтях **СМЕРІЧЧЯ**, **СМЕРЕЧЧЯ**, я, с., збірн.; **ВКАЗІВКА** (**УКАЗІВКА**), и, ж.; **АВАНТЮРА**, рідко **АВАНТУРА**, и, ж.; **БАНКРУТСТВО**, **БАНКРОТСТВО**, а, с.; **КЕТЯГ**, рідко **КИТЯГ**, а, ч. та ін. реалізуються два компоненти, які виділяються в одному числовому комплексі, одному рід-блоці та одному парадигматичному підблоці.

У словникових статтях СЕМИРЯЖКА, СЕМЕРЯЖКА, СЕМРЯЖКА, и, ж., заст.; СЕСТРІНИЦЯ, СЕСТРИНИЦЯ, СЕСТРІНИЦЯ і, ж., діал.; БЕСКІД, БЕСКЕД, БЕСКЕТ, у, ч.; ВНУЧА (ОНУЧА, рідко УНУЧА), ати, с. реалізуються по три варіантних компоненти в межах комплексу, блоку та підблоку.

Таким чином, представлення граматичної семантики іменників у СУМ-11 відбувається відповідно до їх багаторівневої ієрархічної структури, що дозволяє твердити про специфіку якісного й кількісного розподілу структурних елементів і дає змогу сформулювати систему комплексних правил, урахування яких структурує граматичну параметризацію іменників. Виведені правила кількісного співвідношення параметрів граматичної семантики іменників української мови можна представити узагальнено за допомогою графа (див. рис. 3.2).

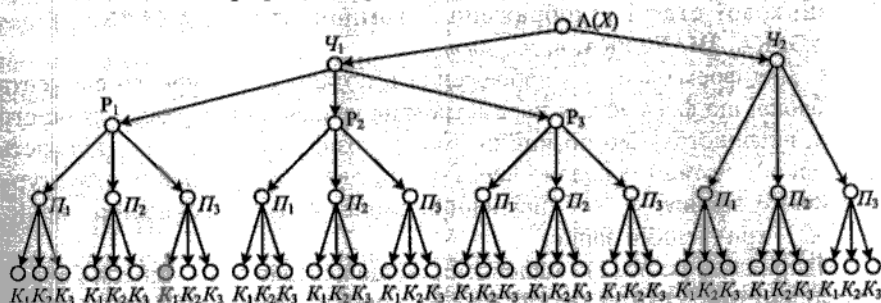


Рис. 3.2. Графічне відображення структури $\Lambda(X)$ іменників української мови у СУМ-11.

Зауважимо, що граф відображає максимальну кількість теоретично можливих структурних елементів на кожному з рівнів граматичної параметризації.

Кожний з виявлених граматичних параметрів характеризується різним кількісним співвідношенням з іншими граматичними параметрами. Специфіка такого співвідношення виявляє особливості граматичної структури досліджуваних іменників. Відтак постає завдання встановлення наявних типів співвідношення, комбінування граматичних параметрів у межах лівих частин словникових статей іменників. Таке завдання вирішувалося в автоматичному режимі. Отримані результати опрацьовано відповідно до виробленої нотації і представлено у вигляді сигнатур⁵ — системи

⁵ Запис сигнатури відображається у круглих дужках ().

аналітично-числового формального представлення структур лівих частин іменникових статей.

Розроблена нотація ґрунтується на системі умовних позначень, окремих для кожного рівня параметризації граматичної семантики іменників української мови в СУМ-11:

1. Для формального відображення елемента «компонент» використовуємо цифрові позначення, які вказують на загальну кількість компонентів, що реалізуються в межах кожного підблоку, блоку та комплексу лівої частини статті. Наприклад, запис (1) вказує на те, що структура лівої частини словникової статті формується на основі одного компонента – заголовного слова: **ПОБІГЬНЬКИ**, *ньюк, мн., розм.* Запис (12) вказує на те, що в структурі статті відображено окремо 1 та 2 компоненти: **САЛГАН**, *а, ч., діал.* Запис (3) вказує на те, що в структурі словникової статті відображено 3 компоненти: **ЗНАХАРКА**, *рідко* **ЗНАХОРКА**, **ЗНАХУРКА**, *и, ж., заст.*

2. Для формального відображення в нотації структурного елемента «підблок», що формується параметром на позначення парадигматичного класу, використовуємо позиційні умовні позначення. Загальну кількість підблоків в нотації передано через одно-, дво- або тричленну структуру числового запису: X_1 – одночленний запис, X_1X_2 – двочленний запис, $X_1X_2X_3$ – тричленний запис.

Наприклад, запис (2) – одночленний, що передає один підблок з двома компонентами: **УГІДДЯ (ВГІДДЯ)**, *я, с.* Запис (3) – також одночленний, він відбиває один підблок з трьома компонентами: **СЕМІРЯЖКА**, **СЕМЕРЯЖКА**, **СЕМРЯЖКА**, *и, ж., заст.* Запис (21) – двочленний, у ньому відображено два блоки з двома та одним компонентами.

3. Для формального відображення в нотації структурного елемента «блок», що формується параметром на позначення роду, застосовуємо позначення – крапку «.» У випадку одночасної наявності двох чи трьох блоків у межах словникової статті цей маркер вказує на їх розмежування. Формально крапка відділяє один блок від іншого. Запис з двома родовими блоками набуває вигляду: P_1P_2 , де P_1 та P_2 – два різних блоки, значення яких стосуються окремих компонентів.

Наприклад, запис (2.1) вказує на те, що ліва частина словникової статті формується на основі двох блоків: 2 та 1, в яких по одному парадигматичному підблоку, в межах яких виділяються, відповідно, два та один компоненти: **МАКОТЕРТЬ**, **МАКОРТЕТЬ**, *і, ж.,* **МАКОРТЕТ**, *а, ч., діал.*

Досі розглядалися приклади нотації формальних структур, у яких до складу блоків та підблоків входили різні компоненти. Проте в СУМ-11

зафіксовані й такі структури, у яких один компонент виявляється спільним для двох підблоків. У такому випадку, в систему цифрової нотації вводимо позначення рівності за компонентом («=k»), яке відбиває спільний компонент для кількох підблоків.

Наприклад: запис (2.=k2) вказує на те, що структура граматичної семантики іменникової лексеми розкладається на два блоки, два підблоки, які об'єднуються двома спільними компонентами: **ДРОЖ**, *розм. ДРІЖ*, *дрожу, ч. і дрóжі, ж.* Запис (1.=k1) вказує на те, що структуру граматичної семантики іменникової лексеми формують один блок, два підблоки, які об'єднуються одним спільним компонентом: **ЖУЖІЛЬ**, *жўжелі, ж. і жўжелю, ч.; ЛАЗЎР*, *і, ж., рідко у, ч.*

Зафіксовані також випадки, коли компонент чи компоненти, що належать до одного парадигматичного підблоку, розмежовуються за різними блоками. Інакше кажучи, компоненти лівої частини словникової статті, що належать до одного парадигматичного класу, можуть набувати двох родових значень. Для відображення таких структур вводимо позначення «.=» – рівність компонентів і підблоків за виділюваними блоками.

Наприклад: запис (1.=1) вказує на те, що в статті відображено один компонент у межах одного підблоку, що входять до складу двох різних блоків: **ПУТЬ**, *і, ж., заст. ч.; СОНЬКО, *а, ч. і ж., розм.* Запис (3.=3) означає, що в структурі лівої частини словникової статті – три компоненти, що входять до одного підблоку й розмежовуються за двома різними блоками: **БІДОЛА́ХА**, *рідко БІДОЛА́ГА, БІДОЛА́КА*, *и, ч. і ж., розм.* Запис (1.=1.=1) вказує на те, що в структурі лівої частини словникової статті міститься один компонент, що належить одному підблоку, але він розмежовується за трьома блоками: **ЛА́ДО**, *а, ч., ж. і с. нар.-поет.**

4. Згідно з числовою параметризацією, ліві частини словникових статей іменників можуть бути одно- або двокомплексними.

Для відображення подвійних числових значень числового комплексу застосовуємо позначення «:» як спосіб розмежування комплексів у межах лівої частини словникової статті. Відтак відображення формальних показників, що стосуються окремо однини та множини, здійснюється через ліво- та правосторонній запис щодо «:». Наприклад: словникова стаття **ТУАРЕ́ГИ**, *гів, мн. (одн. ТУАРЕ́Г, а, ч.)* відображається за допомогою аналітично-числового запису (1:1).

Нагадаємо, що до двокомплексних відносимо також іменники з повною і нормативною числовою парадигмою, яка, однак, у СУМ-11, згідно з вимогами лексикографування, передається частково: фіксуються лише

граматичні параметри в однині, при цьому без числового ремаркування:
ПАСАТ, у, ч.; **ПАСКА**, и, ж.; **ПАСТОРСТВО**, а, с.

Цей лексикографічний факт важливо відзначити, оскільки в нашому дослідженні встановлення структури граматичної семантики іменників здійснювалося автоматично за наявності або відсутності відображуваних у лівій частині словникової статті параметрів числа: *одн., мн.* або *збірн.* За фактичної відсутності у словниковій статті такого типу числового показника самостійно визначаємо структуру граматичної семантики цього класу іменників як двочислової і присвоюємо їй значення сигнатури (1:). Відсутність числа після двокрапки вказує на наявність множини, що припускається у мові. До двочислових, але з умовно наявною множиною відносимо статті: **САЛГАН**, а, **САЛОГАН**, **САЛОГОН**, а, ч., *діал.* → (12:); **ДЕБРА**, и, *рідко* **ДЕБР**, **ДЕБРЬ**, і, ж., *зах.* → (12:); **ГЛЯНЕЦЬ**, и, що, **ГЛЯНС**, **ГЛЯНЦ**, у, ч. → (12:). Їх структура визначається формальним записом (12:), де форма множини припускається як об'єктивно існуюча, хоч вона й не відображена у СУМ-11.

У словникових статтях одноконлексні структури іменників, утворювані неповною числовою парадигмою (класи збірних іменників, *singularia tantum* і *pluralia tantum*), що маркуються у СУМ-11 як *збірн., мн.* або *тільки одн.*, відображаємо за допомогою умовного позначення латинською літерою *n*:

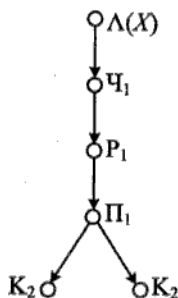
СМЕРІЧЧЯ, **СМЕРЕЧЧЯ**, я, с., *збірн.*, де $Ч_1 =$ **СМЕРІЧЧЯ**, **СМЕРЕЧЧЯ**, я, с., *збірн.* Набуває аналітично-числового запису (2n); **ОШКРЄБКИ**, **ОСКРЄБКИ**, ів, мн., *рідко*, де $Ч_1 =$ **ОШКРЄБКИ**, **ОСКРЄБКИ**, ів, мн., *рідко*. → (2n); **ЛЬОН**, у, ч., *тільки одн.*, де $Ч_1 =$ **ЛЬОН**, у, ч., *тільки одн.* → (1n).

У двоконлексних структурах умовне позначення *n* фіксуємо лише у препозиції щодо комплексу зі значенням однини. Воно вказує на число множинне або збірне. Наприклад: **ПІДСЛУХИ**, ів, мн. (*одн.* **ПІДСЛУХ**, у, ч.) → (1n:1) (маркується перший числовий комплекс); **ХАЗЯЇН**, а, ч. (*мн.* **ХАЗЯЇНИ**, ів, і **ХАЗЯЇ**, ів) → (1:1) (другий числовий комплекс не маркується).

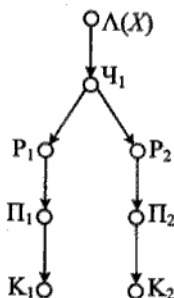
Згідно з описаною нотацією, словникова стаття **ПЕЙЗАНИ**, ів, мн. (*одн.* **ПЕЙЗАН**, **ПЕЙЗАНИН**, а, ч.; **ПЕЙЗАНКА**, и, ж.) одержує формальний запис (1n:2.1). Він вказує на те, що структура статті має два комплекси (про що свідчить позначення «:»), перший з яких характеризується маркованим числовим значенням (множинне), складається з одного підблоку та одного компонента; другий числовий комплекс формується на основі двох блоків (ч. і ж.), перший з яких має один підблок і два компоненти, а другий блок містить один підблок і один компонент.

Отже, ліву частину іменникових статей СУМ-11 можна представити за допомогою системи сигнатур та в графічному зображенні. Наведемо приклади деяких з них:

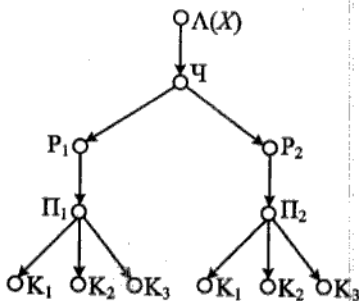
МІЛЬЙОНЕРКА, *заст. МІЛОНЕРКА*, *и, ж., розм.* → (2:).



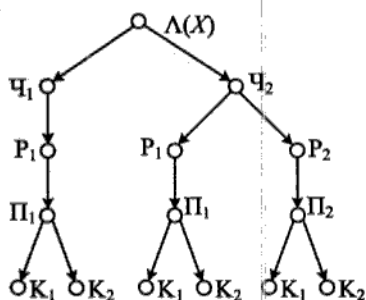
ВУАЛЬ, *і, ж., рідко я, ч.* → (1.=k1).



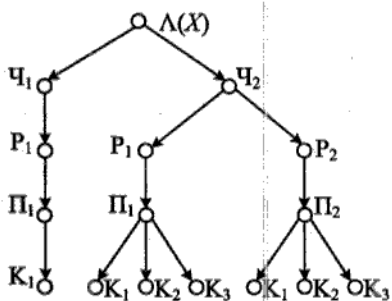
СЕРДЕГА, *рідко СЕРДЕКА, СЕРДЯГА*, *и, ч. і рідше ж.* → (3.=3).



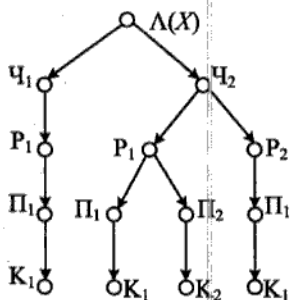
ОДНОСЕЛЬЧАНИ, ОДНОСІЛЬЧАНИ, чан, мн. (одн. **ОДНОСЕЛЬЧАНИН, ОДНОСІЛЬЧАНИН**, а, ч.; **ОДНОСЕЛЬЧАНКА, ОДНОСІЛЬЧАНКА**, и, ж.) → (2п:2.2).



СЕРБИ, ів, мн. (одн. **СЕРБ**, рідше **СЕРБИН**, заст. **СЕРБІЯНИН**, а, ч.; **СЕРБКА**, рідше **СЕРБИНКА**, заст. **СЕРБІЯНКА**, и, ж.) → (1п:3.3).



ТАТАРИ, ів, мн. (одн. **ТАТАРИН**, а, рідко **ТАТАР**, а, ч.; **ТАТАРКА**, и, ж.) → (1п:11.1).



За допомогою послідовного застосування теорії лексикографічних систем встановлено параметризацію граматичної семантики іменників української мови, відображених у СУМ-11. Розподіл іменників за сигнатурою дозволив здійснити детальний аналіз виділених класів з різним співвідношенням структурних елементів (див. табл. 3.8).

Таблиця 3.8. Розподіл іменників сучасної української мови за сигнатурами

№	Сигнатура	Загальна кількість	Іменникова стаття Л(Х)
1.	(:)	50041	ПІВУСТАВ, у, ч.
2.	(1n)	1083	ЧЕРЕПИЦЯ, і, ж., збірн.
3.	(2:)	873	ПРОЗОРИВІСТЬ, рідко ПРОЗІРЛИВІСТЬ, вості, жс.
4.	(1=1:)	365	СЕМІЛІТОК, тка, ч. і жс., розм.
5.	(1n:1.1)	306	СВАНИ, ів, мн. (одн. СВАН, а, ч.; СВАНКА, и, жс.).
6.	(1n:1)	70	ГАЛЛИ, ів, мн. (одн. ГАЛЛ, а, ч.)
7.	(11:)	49	СРІБЛО-ЗОЛОТО, сріблА-зОлота, СРІБЛО-ЗЛОТО, сріблА-злАта, с., нар.-поет.
8.	(1.1:)	28	РУКОПИС, у, ч., розм., рідко РУКОПИСЬ, і, жс.
9.	(1≠k1:)	19	ЧЕРІНЬ, реня, ч. і рені, жс.
10.	(3:)	18	СЕМІРЯГА, СЕМЕРЯГА, СЕМРЯГА, и, жс., заст.
11.	(2n)	17	УСТА (ВУСТА), уст, мн.
12.	(11n)	10	ПЕРЛИ, ів, рідко ПЕРЛА, рел. мн.
13.	(2.=2:)	10	РОЗУМАКА, РОЗУМАХА, и, ч. і жс., розм.
14.	(12:)	4	САЛГАН, а, САЛОГАН, САЛОГОН, а, ч., діал.
15.	(2n:2.2)	4	ОСТРІВ'ЯНИ, ОСТРОВ'ЯНИ, ян, мн. (одн. ОСТРІВ'ЯНИН, ОСТРОВ'ЯНИН, а, ч.; ОСТРІВ'ЯНКА, ОСТРОВ'ЯНКА, и, жс.).
16.	(1n.11.1)	3	ТАТАРИ, ів, мн. (одн. ТАТАРИН, а, рідко татар, а, ч.; ТАТАРКА, и, жс.).
17.	(2n.11.1)	3	СКАНДІНАВИ, СКАНДІНАВЦІ, ів, мн. (одн. СКАНДІНАВ, а, СКАНДІНАВЕЦЬ, вия, ч.; СКАНДІНАВКА, и, жс.).
18.	(1=1=1:)	2	ЛАДО, а, ч., жс. і с., нар.-поет.
19.	(1n1=1)	2	УДЕ, мн. (одн. УДЕ, ч. і жс.), невідм.
20.	(1n.1.11)	2	ГРЕКИ, ів, мн. (одн. ГРЕК, а, ч.; ГРЕЧАНКА, и і ГРЕКИНЯ, і, жс.).
21.	(1n:1.2)	2	ПОЛЯКИ, ів, мн. (одн. ПОЛЯК, а, ч.; ПОЛЬКА, заст. ПОЛЯЧКА, и, жс.).
22.	(2.=k2:)	2	ДРОЖ, розм. ДРІЖ, дрОжу, ч. і дрОжі, жс.
23.	(3=3:)	2	СЕРДЕГА, рідко СЕРДЕКА, СЕРДЯГА, и, ч. і рідше жс.
24.	(111:)	1	КАДЬ, доба, КАДІВЬ, довба, рідко КАДУБ, а, ч.
25.	(11n:1.1)	1	САБІНИ, ів, мн. (одн. САБІНЯНИ, нян, мн. (одн. САБІНЯНИН, а, ч.; САБІНЯНКА, и, жс.).
26.	(1n:2.1)	1	ПЕЙЗАНИ, ів, мн. (одн. ПЕЙЗАН, ПЕЙЗАНИН, а, ч.; ПЕЙЗАНКА, и, жс.).
27.	(1n:3.3)	1	СЕРБИ, ів, мн. (одн. СЕРБ, рідше СЕРБИН, заст. СЕРБІЯНИН, а, ч.; СЕРБКА, рідше СЕРБИНКА, заст. СЕРБІЯНКА, и, жс.).
28.	(2.1:)	1	МАКОТЕРТЬ, МАКОРТЕТЬ, і, жс., МАКОРТЕТ, а, ч., діал.
29.	(21:)	1	ЯТЕЛ, ЯТІЛ, ятла, ЯТІЛЬ, ятєля, ч., діал.
30.	(2n:2)	1	ШКАРБАНИ, ШКАРБУНИ, ів, мн. (одн. ШКАРБАН, ШКАРБУН, а, ч.), розм.
31.	(3n)	1	ВИШКРЕБКИ, ВИСКРЕБКИ, ів, мн., розм.

Представлена формальна класифікація іменників за параметрами їх граматичної семантики дозволяє проаналізувати відповідні статистичні дані й встановити та верифікувати певні закономірності щодо цього класу слів у системі української мови.

Як показав аналіз числової параметризації іменників української мови, їх можна розділити за трьома класами:

1) іменники, граматична семантика яких характеризується наявністю корелятивних числових форм однини та множини. Зазначимо, що за правилами лексикографування у СУМ-11 такі лексеми описуються через фіксування граматичних параметрів лише в однині, при цьому їх множинні форми припускаються за умовчанням;

2) іменники, структура граматичної семантики яких визначена лексикографічним способом їх опрацювання (назви народностей);

3) іменники, які не мають корелятивної форми множини (*Singularia tantum* та збірні) чи однини (*Pluralia tantum*) (див. табл. 3.9).

Таблиця 3.9. Розподіл іменників за параметром числа

№	Іменниковий клас	Кількість	%
1.	Іменники з корелятивними формами однини та множини	51416	97,1
2.	Іменники, які не мають корелятивної форми множини (<i>Singularia tantum</i> та збірні) чи (<i>Pluralia tantum</i>)	1111	2

Аналіз кількісних даних, наведених у таблиці, свідчить про те, що в українській мові граматична категорія числа дуже розвинута. Перевага іменникових лексем з корелятивними формами числа становить 97,1% від загальної кількості іменників, лише 2% припадає на іменники, які не мають однієї з корелятивних форм (одночислові).

Як окремо виділений, згідно з правилами лексикографування, розглянемо клас іменників на позначення назв народностей.

Найчисельнішими і найтиповішими тут виявляються лексеми – назви народностей, структура граматичної семантики яких передбачає відображення в числовому комплексі однини двох параметрів роду – чоловічого й жіночого, кожний з яких співвідноситься з одним компонентом: ГОЛЛАНДИ, ів, мн. (одн. ГОЛЛАНДЕЦЬ, дця, ч.; ГОЛЛАНДКА, и, ж.); ВОЛИНЯКИ, ів, мн. (одн. ВОЛИНЯК, а, ч.; ВОЛИНЯЧКА, и, ж.), розм.; ВІННИЧАНИ, ан, мн. (одн. ВІННИЧАНИН, а, ч.; ВІННИЧАНКА, и, ж.) (див. табл. 3.10).

Таблиця 3.10. Розподіл іменникових лексем на позначення назв народностей за сигнатурами

№	Назви народностей	Тип сигнатури	Кількість лексем у межах сигнатури
1.	МЕЩЕРЯКИ, ів, мн. (одн. МЕЩЕРЯК, а́, ч.; МЕЩЕРЯЧКА, и, жс.).	(1n:1.1)	306
2.	САБІНИ, ів, САБІНЯНИ, нян, мн. (одн. САБІНЯНИН, а, ч.; САБІНЯНКА, и, жс.).	(11n:1.1)	1
3.	ПЕЙЗАНИ, ів, мн. (одн. ПЕЙЗАН, ПЕЙЗАНИН, а, ч.; ПЕЙЗАНКА, и, жс.).	(1n:2.1)	1
4.	КУРДИ, ів, мн. (одн. КУРД, а, ч.; КУРДКА і КУРДЯНКА, и, жс.).	(1n:1.2)	2
5.	ТАТАРИ, ів, мн. (одн. ТАТАРИН, а, рідко ТАТАР, а, ч.; ТАТАРКА, и, жс.).	(1n:11.1)	3
6.	НІМЦІ, ів, мн. (одн. НІМЕЦЬ, мця, ч.; НІМКЕНЯ і НІМКА, и, жс.).	(1n:1.11)	2
7.	УДІ, мн. (одн. УДІ, ч. і жс.), невідм.	(1n:1.=1)	2
8.	СЕРБИ, ів, мн. (одн. СЕРБ, рідше СЕРБИН, заст. СЕРБЯНИН, а, ч.; СЕРБКА, рідше СЕРБИНКА, заст. СЕРБЯНКА, и, жс.).	(1n:3.3)	1
9.	БАЛКАРИ, БАЛКАРЦІ, ів, мн. (одн. БАЛКАР, а і БАЛКАРЕЦЬ, рця, ч.; БАЛКАРКА, и, жс.).	(2n:11.1)	3
10.	ІУДЕЇ (ЮДЕЇ), ів, мн. (одн. ІУДЕЙ, ЮДЕЙ, я, ч.; ІУДЕЙКА, ЮДЕЙКА, и, жс.).	(2n:2.2)	4
11.	КОНКІСТАДОРИ, КОНКВІСТАДОРИ, ів, мн. (одн. КОНКІСТАДОР, КОНКВІСТАДОР, а, ч.).	(2n:2)	1
12.	ГОЛЬДИ, ів, мн. (одн. ГОЛЬД, а, ч.), заст..	(1n:1)	70
Всього			396

Як засвідчують сигнатури (1n:2.1) та (1n:1.2), кількість варіантних компонентів при одному роді може сягати максимально двох. Такою структурою в українській мові характеризується лише три лексеми на позначення назв народностей (див. табл. 3.11).

Таблиця 3.11. Компонентна варіативність у структурах граматичної семантики іменників на позначення народностей

№	Тип сигнатури	Лексеми – назви народностей
1.	(1n:2.1)	ПЕЙЗАНИ, ів, мн. (одн. ПЕЙЗАН, ПЕЙЗАНИН, а, ч.; ПЕЙЗАНКА, и, жс.).
2.	(1n:1.2)	КУРДИ, ів, мн. (одн. КУРД, а, ч.; КУРДКА, КУРДЯНКА, и, жс.).
3.		ПОЛЯКИ, ів, мн. (одн. ПОЛЯК, а, ч.; ПОЛЬКА, заст.. ПОЛЯЧКА, и, жс.).

Малочисельними виявляються структури граматичної семантики іменників – назв народностей з ускладненням сигнатури на рівні словозміни і компонентної лексемної варіації. Так, унікальною за структурою граматичної семантики виявляється лексема **САБІНИ** – сигнатура (1In:1.1), яка в українській мові має абсолютний синонім у множині, що визначається іншим парадигматичним класом: **САБІНИ**, ів, **САБІНЯНИ**, нян, *мн.* (одн. **САБІНЯНИН**, а, ч.; **САБІНЯНКА**, и, ж.).

Малочисельними є іменники на позначення назв народностей, граматична семантика яких ускладнена наявністю варіативних компонентів як абсолютних синонімів, що належать до різних парадигматичних класів у чоловічому чи жіночому роді. Наприклад: сигнатура (1n:11.1): **ТАТАРИ**, ів, *мн.* (одн. **ТАТАРИН**, а, рідко **ТАТАР**, а, ч.; **ТАТАРКА**, и, ж.). Сигнатура (1n:1.11): **НІМЦІ**, ів, *мн.* (одн. **НІМЕЦЬ**, мця, ч.; **НІМКЕНЯ**, і **НІМКА**, и, ж.); **ГРЕКИ**, ів, *мн.* (одн. **ГРЕК**, а, ч.; **ГРЕЧАНКА**, и і **ГРЕКІНЯ**, і, ж.).

Досі розглядалися типи сигнатур з корелятивними значеннями роду (ч. і ж.). Однак, як засвідчує сигнатура (1n:1), частина лексем на позначення назв народностей характеризується відсутністю лексемних відповідників на визначення форми жіночого роду. До них належать іменники типу **АВАРИ**, ів, *мн.* (одн. **АВАР**, а, ч.); **АЦТЄКИ**, ів, *мн.* (одн. **АЦТЄК**, а, ч.); **БАСКИ**, ів, *мн.* (одн. **БАСК**, а, ч.); **БЛОПОЛЯКИ**, ів, *мн.* (одн. **БЛОПОЛЯК**, а, ч.), *іст.*; **БУРИ**, ів, *мн.* (одн. **БУР**, а, ч.) та ін.

За наявністю числового параметра у структурі граматичної семантики іменників виділяються групи з повною і неповною числовими парадигмами. При цьому у групі іменникових лексем, які не мають однієї з корелятивних форм, встановлюється певний розподіл за конкретним параметром числа (див. табл. 3.12).

Таблиця 3.12. Розподіл іменників з неповною числовою парадигмою

№	Іменниковий клас	Кількість	%
1.	Іменники <i>Singularia tantum</i> ⁶	8	0,01
2.	Іменники збірні	253	0,5
3.	Іменники <i>Pluralia tantum</i>	850	1,6

⁶ У даному випадку йдеться лише про ті іменники, які в СУМ-11 маркуються за допомогою позначення *тільки одн.* Значна кількість цих іменників у СУМ-11 не має ніяких умовних позначень, тому виділити їх за формальними лексикографічними параметрами неможливо.

Іменники *Pluralia tantum* з погляду семантики виявляють співвідно-
ність з граматичним значенням форм однини і множини, тому вони вира-
жають омонімію обох чисел.

Аналіз сигнатур за параметром на позначення роду виявив цікаве кіль-
кісне співвідношення іменників української мови за родом (див. табл. 3.13).

Таблиця 3.13. Розподіл іменників за параметром роду

№	Іменниковий клас	Кількість			%		
		ч.	ж.	с.			
1.	Іменники з одним родовим зна- ченням (ч., або ж., або с.)	19660	19597	10316	37,1	37	19,5
2.	Іменники з двома параметрами роду	427			0,8		
3.	Іменники з трьома параметрами роду	2			0,004		

Окрему групу в українській мові становлять іменники, які в одній і
тій самій формі можуть вживатися із значенням обох родів – чоловічого
й жіночого, чоловічого й середнього та жіночого й середнього. Це
іменники подвійного роду при одному компоненті: **ПРИБЛУДА**, и, ч. і
ж., *зневажл.*; **СТРАШІЛО**, а, ч. і с., *розм.*; **СТИЛЯГА**, и, ч. і ж., *розм.*;
СОБАЧКА, и, ч. і рідше ж.; **ОДНОРУЧКА**, и, ж. і ч., *рідко.*; **РОЗУМНИЦЯ**,
і, ж. і рідко ч., *розм.*; **АГАКАЛО**, а, с. і ч., *зневажл.*; **ПЕННИ**, невідм., с., *рідко*
ч.; **КОНТРАЛЬТО**, невідм., с. і ч.; подвійного роду при двох компонентах:
РОЗУМАКА, **РОЗУМАХА**, и, ч. і ж., *розм.*; **БІДАХА**, *рідко* **БІДАГА**, и, ч. і
ж.; **ЗЛОДНЮГА**, **ЗЛОДНЮКА**, и, ч. і ж., *розм.*; **СКУПЕНДРА**, **СКУПІНДРА**,
и, ч. і ж., *розм.*; подвійного роду при трьох варіативних компонентах:
СВЕКРУШИЩЕ, а, с. і ж., *зневажл.*; **СЕРДЕГА**, *рідко* **СЕРДЕКА**,
СЕРДЯГА, и, ч. і рідше ж.; **БІДОЛАХА**, *рідко* **БІДОЛАГА**, **БІДОЛАКА**, и, ч. і
ж., *розм.*

За допомогою автоматичного аналізу структури граматичної семанти-
ки іменників виявляються також лексеми з потрійним родом, причому з
різним комбінуванням параметрів роду: **ЛАДО**, а, ч., ж. і с., *нар.-поет.*;
ЗАБУДЬКО, а, ч., с. і ж., *розм.*

Окремо виділяється клас іменників з різним граматичним оформлен-
ням, що становлять відмінні парадигматичні системи, належать до різних
словозмінних класів, але характеризуються спільним родом. З такою
структурою у СУМ-11 зафіксовано п'ять лексем: **САЛГАН**, а, **САЛОГАН**,
САЛОГОН, а, ч., *діал.*; **ЗНАХАР**, я, *рідко* **ЗНАХОР**, **ЗНАХУР**, а, ч., *заст.*;
ДЕБРА, и, *рідко* **ДЕБР**, **ДЕБРЬ**, і, ж., *зах.*; **ГЛЯНЕЦЬ**, ищо, **ГЛЯНС**, **ГЛЯНЦ**, у,
ч., **ЯТЕЛ**, **ЯТІЛ**, ятла, **ЯТІЛЬ**, ятеля, ч., *діал.*

Відповідно до правила 3 побудови структури граматичної семантики іменників української мови, варіювання словозмінних класів, що стосуються реєстрового слова, максимально дорівнює трьом. Відтак, у межах одного роду виділяємо іменники, що визначаються: 1) одним парадигматичним класом – 52032 (98,3%); 2) двома парадигматичними класами – 144 (0,3%); 3) трьома парадигматичними класами – 1 (0,001%).

Виходячи з цих даних, можна зробити висновок про те, що словозмінна розгалуженість для української мови є нерегулярною. На це вказує незначна кількість іменників, що в своїй структурі граматичного значення мають таке розгалуження за двома словозмінними параметрами, а структура з трьома парадигматичними класами виявляється взагалі унікальною: **КАДІБ**, доба, **КАДІВБ**, довба, *рідко* **КАДУБ**, а, ч.

Аналіз лексем за сигнатурами дозволяє також виявити клас іменників, які визначаються однією парадигмою в межах двох або трьох параметрів роду. Це так звані дво- і триродові іменники (див. табл. 3.14).

Таблиця 3.14. Розподіл іменників за сигнатурами відносно кількості компонентів у $\Lambda(X)$

№	Іменниковий клас	Кількість компонентів	Кількість лексем	%
1.	Іменники з двома параметрами роду і спільним парадигматичним класом	1	367	0,7
		2	10	0,02
		3	2	0,003
2.	Іменники з трьома параметрами роду і спільним парадигматичним класом	1	2	0,003
		2	–	–
		3	–	–

Окрему увагу звернімо на клас іменникових лексем, структура граматичної семантики яких реалізується сигнатурами (1= k_1) та (2= k_2): **КУ-ПІЛЬ**, купели *ж. і купелю ч.*; **ЖУЖІЛЬ**, жужелі, *ж. і жужелю ч.*; **ХАРЧ**, у, ч. і і, *ж.*; **ЧЕРІНЬ**, реня, ч. і рені, *ж.*; **ОПАР**, у, ч., *рідко і, ж., діал.*; **ШАШІЛЬ**, шеля, ч. і розм. шелі, *ж.*; **ФАЛЬШ**, і, *ж. і у, ч.*; **ПЕРЕСІП**, у, ч. і діал. і, *ж.*; **ПÓКРУЧ**, *пóкруча, ч., пóкручі, ж., розм.*; **ШИР**, у, ч. і і, *ж., рідко*; **ДРОЖ**, *розм. ДРІЖ*, *дрóжу, ч. і дрóжі, ж.*; **ÓВОЧ**, *заст. ÓВОЦ*, у, ч. і *рідко і, ж.*

Аналіз структури граматичної семантики таких лексем свідчить, що вони є прикладом граматичної омонімії щодо спільного компонента чи компонентів, які стосуються двох властивих їм параметрів роду, в кожному з яких він набуває окремої словозмінної характеристики. В українській мові зафіксовано 21 іменник з такою граматичною структурою: з них

19 лексем визначаються сигнатурою (1.=k1:): **КУПІЛЬ**, купелі, ж. і купелю, ч.; **ЧОРНОБИЛЬ**, лю, ч. і розм. лі, ж. (*Artemisia vulgaris* L) та дві лексеми сигнатурою (2=k2:): **ДРОЖ**, розм. **ДРІЖ**, дрóжу, ч. і дрóжі, ж.; **ОВОЧ**, заст. **ОВОЩ**, у, ч. і рідко і, ж.).

Лексикографічний опис таких слів демонструє згортання всієї інформації у таких лексемах. Структуру їх граматичної семантики можна передати за допомогою розгорнутого запису: **КУПІЛЬ**, купелі, ж. і купелю, ч.; **КУПІЛЬ**, купелі, ж. і **КУПІЛЬ**, купелю, ч.; **ЧОРНОБИЛЬ**, лю, ч. і розм. лі, ж. (*Artemisia vulgaris* L); **ЧОРНОБИЛЬ**, лю, ч. і **ЧОРНОБИЛЬ**, розм. лі, ж. (*Artemisia vulgaris* L); **ДРОЖ**, розм. **ДРІЖ**, дрóжу, ч. і дрóжі, ж.; **ДРОЖ**, розм. **ДРІЖ**, дрóжу, ч. і **ДРОЖ**, розм. **ДРІЖ**, дрóжі, ж.; **ОВОЧ**, заст. **ОВОЩ**, у, ч. і рідко і, ж.; **ОВОЧ**, заст. **ОВОЩ**, у, ч. і **ОВОЧ**, заст. **ОВОЩ**, рідко і, ж.

Як уже зазначалося, компонентний склад структури граматичної семантики іменників української мови формується на основі префіксальної, суфіксальної, префіксально-суфіксальної та коренево-суфіксальної лексемної варіації реестрового слова. При цьому компоненти, що визначаються однаковими граматичними параметрами словозміни, роду й числа, максимально мають три варіанти. Вживання таких варіантів зумовлене функціональними (стильовими) особливостями. Виходячи із встановленої можливої кількості компонентів у межах однієї граматичної параметризації виділяємо групи іменників: 1) з одним компонентом – 50041 (94,5%); 2) двома компонентами – 873 (1,6%); 3) трьома компонентами – 18 (0,03%).

Іменники з трьома компонентами визначаються кількома сигнатурами, деякі з них є унікальними (див. табл. 3.15).

Таблиця 3.15. Розподіл іменників за наявністю лексемного варіанта в структурі їх граматичної семантики

№	Сигнатура	Іменникова лексема	Кількість
1.	(3:)	ДЗБАН, ДЖБАН , розм. ЖБАН , а, ч.	18
2.	(3.=3:)	СЕРДЕГА , рідко СЕРДЕКА, СЕРДЯГА , и, ч. і рідше ж. БІДОЛАХА , рідко БІДОЛАГА, БІДОЛАКА , и, ч. і ж., розм.	2
3.	(3n)	ВІШКРЕБКИ, ВІСКРЕБКИ, ВІСКРІБКИ , ів, мн., розм.	1
4.	(1n:3.3)	СЕРБИ , ів, мн. (одн. СЕРЬ , рідше СЕРБИН , заст. СЕРБІЯНИН , а, ч.; СЕРБКА , рідше СЕРБИНКА , заст. СЕРБІЯНКА , и, ж.).	1

Аналіз матеріалу СУМ-11 показує, що компонентні лексемні варіації не є регулярними в українській мові, а поява додаткового компонента веде до різкого та багатократного скорочення кількості лексем з такою граматичною структурою.

Таким чином, проведене дослідження дозволяє твердити про існування нового класифікаційного критерію, на основі якого може бути верифіковано та систематизовано особливості функціонування граматичної підсистеми іменників у системі української мови. Необхідну для цього інформацію, виявлену за допомогою застосування методологічних засад теорії лексикографічних систем, представлено в текстах тлумачних словників системно й упорядковано, що зумовлює її адекватну формалізацію і дозволяє провести класифікацію іменників української мови за їх формальними параметрами.

3.4. Мікроструктура інтерпретаційних частин словникових статей СУМа

У цьому підрозділі розглядається структура $P(X)$ як репрезентанта змісту мовного знака в СУМі, де у правих частинах міститься детальна семантична інформація про реєстрові одиниці.

Праві (інтерпретаційні) частини словникових статей СУМа за своїми розмірами, як правило, набагато ширші від відповідних реєстрових (лівих) частин. Це зумовлюється функціонально-інтерпретаційною орієнтацією цього елемента словникової статті. Іншими словами, праву частину можна розглядати як структурований певним чином та проінтерпретований фрагмент лексичної картотеки або лінгвістичного корпусу. Якщо множина лівих частин словникових статей СУМа репрезентує систему граматичної семантики української мови, то відповідна множина правих частин його словникових статей репрезентує систему лексичної семантики. Отже, її обсяг тлумачної частини часто в десятки разів перевищує обсяг лівої частини. Водночас структурна організація правої частини у СУМі є певною мірою простішою. Не заглиблюючись в аналіз та тлумачення цього факту, перейдемо до опису структуротвірних елементів $P(X)$.

За аналогією до розгляду лівої частини, досліджуючи будову $P(X)$, виходимо з того, що права частина також складається з декількох вкладених одна в одну підструктур, ідентифікація яких пов'язана з певними лінгвістичними феноменами.

У структурі правої частини $P(X)$ словникової статті СУМа відображаються ефекти багатозначності реєстрової одиниці та її фразеологічна час-

тина. Багатозначність має двоступеневий характер і складається з рубрик, які репрезентують лексичні значення відповідних лексем та їх відтінки. Усі структурні ефекти, які ми тут описуємо, формально визначено в тексті СУМа за допомогою спеціальних розділових позначок, позиційно, шрифтовими виділеннями тощо. Тому побудова формального репрезентанта структури може бути повністю алгоритмізована.

Аналіз структур правих частин $P(X)$ словникових статей СУМа дозволяє виокремити певні структуротвірні елементи. Найбільшою з цих структур є та, яка відтворює явище багатозначності (полісемії) реєстрової одиниці. Отже, полісемія S_i – це частина $P(X)$, яка дає тлумачення i -того значення реєстрової одиниці x . Звідси, будь-яку $P(X)$ можна подати як об'єднання

$$P(X) = \bigcup_{i=1}^{n(x)} S_i, S_i \cap S_j = \emptyset \text{ при } i \neq j. \quad (3.32)$$

У кожній рубриці S_i розміщені ілюстрації i -того значення, тобто приклади його використання в літературному тексті, які позначаємо символом $J(i, q)$, а також фрагменти словникової статті, які позначають її відтінки. У поліграфічному оформленні відтінки представлені двома скісними рисками – //; відтінки значення S_i позначатимемо через $V(i, j)$. До кожного відтінку може подаватися одна або кілька ілюстрацій. Символом $J^v(i, j, k)$ позначимо k -ту ілюстрацію відтінку $V(i, j)$.

Таку саму структуру у СУМ-20 має фразеологічна система⁷ (сукупність різних типів усталених словосполучень), лише з тією різницею, що в ній відсутній поділ на відтінки. У загальному випадку до правої частини $P(X)$ може належати цілий комплекс усталених словосполучень, представлених системами стійких словосполучень, словосполучень-еквівалентів слова, термінологічних словосполучень та власне фразеологічних одиниць. Кожен із цих чотирьох типів усталених словосполучень може мати одне чи декілька значень $FS(i, j, k)$; до кожного значення подаються ілюстрації $J^{FS(i, j, k, m)}$.

З урахуванням викладеного справедливою є система вкладень:

$$\begin{aligned} & J(i, q) \\ & \cap \\ & S_i \supset V(i, r) \supset J^v(i, r, t). \\ & \cup \\ & \{F(i, j) \supset FS(i, j, k) \supset J^{FS(i, j, k, m)}\} \end{aligned} \quad (3.33)$$

⁷ Про подання фразеологічної системи у СУМ-20 детально див. підрозділ 5.5.

Якщо за аналогією з попереднім виконати заміну: $A \supset B \Rightarrow A \rightarrow B$, одержимо загальний вигляд графічної репрезентації структури $P(X)$ СУМ-20 (див. рис. 3.3).

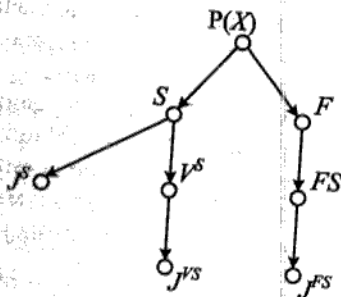


Рис. 3.3. Графічна репрезентація структури правої частини словникової статті у СУМ-20

Якщо в структурних графах деякої словникової статті на відповідних поверхах замість символів подати конкретні значення відповідних атрибутів, то одержимо конструкцію, що є ізоморфною саме цій словниковій статті. Тобто з такого графа цілком однозначно відтворюється словникова стаття у тому ж вигляді, який вона має у відповідному паперовому виданні. Проілюструємо викладене на прикладі словникової статті досить загального вигляду:

БЛАГОСЛОВЛЯТИ, яю, яєш, недок., **БЛАГОСЛОВИТИ**, влію, віш; мн. благословлять; док. і рідко недок., кого, що. 1. бібл. Давати милість, благодать, допомогу і т. ін. кому-, чому-небудь; обдаровувати благами (про Бога). Ти, Господи, благословлятимеш праведного, милістю вкрисис його, як щитом! (Біблія. Пер. І. Огієнка); Нехай благословляє нас Бог, і всі кінці землі хай бояться Його! (Біблія. Пер. І. Огієнка); // кого ким, чим, розм. Давати, дарувати (про Бога). Живе Олеся рік і другий .. і четвертий. Благословив Господь діточками: три сини як три соколи (Марко Вовчок).

2. рел.-церк. Хрестити кого-небудь, проказуючи при цьому молитву, побажання. — Що ж ти нас не благословляєш, панотче? — каже Сомко Череваневі (П. Куліш); Веде Олександра діток, обливаючись сльозами, та все тільки благословляє їх та хрестить... (Марко Вовчок); Сиділа [Згода] на балконі нового будинку та все лиш усміхалася та благословила (І. Франко); Дай вам Боже... дочечку до року до вінця благословити... (Леся Українка); Благослови ж, матусю, чумаків На їх непевну і тяжку дорогу (М. Рильський); Перед початком роботи делегатів та гостей благословили на добро священики Української автокефальної православної та Української греко-католицької церков (з газ.); * Образно. Природа .. лагідно благословляла людські веселощі (Я. Качура).

3. *перев. на що. Давати згоду на що-небудь, дозволяти кому-небудь щось. – Я піду у черницю, – промовляє Катря.. – Я [батько] ніколи тебе на се не благословлю!* (Марко Вовчок); *Вмер старий Варлаам..., перед смертю благословивши свого учня йти в печери* (М. Грушевський).

4. *уроч. Прославляти, хвалити кого-, що-небудь. Я благословлятиму Господа кожного часу, хвала Йому завжди в устах моїх!* (Біблія. Пер. І. Огієнка); *Умийся, серце, щоб пізнала [мати] тебе.. І Господа благословляла за долю добру твою* (Т. Шевченко); *Старий Гурте?.. поселився у капітана і благословить його ім'я та пильнує його дітей* (І. Франко).

5. *Бажати кому-, чому-небудь добра, миру, благодаті і т. ін. Любіть ворогів своїх, благословляйте тих, хто вас проклинає, творить добро тим, хто ненавидить вас, і молитесь за тих, хто вас переслідує* (Біблія. Пер. І. Огієнка); *Не платіть злом за зло або лайкою за лайку, навпаки, благословляйте, знаючи, що на це вас покликано, щоб ви успадкували благословення* (Біблія. Пер. І. Огієнка).

◇ (1) **Божє (Господи) благослові** – уживається як побажання успіху перед початком якої-небудь справи, схвалення якихось дій. – *Благословіть сей чесний і важкий хліб у піч посадити!* – *Божє благослови!* (Г. Квітка-Основ'яненко); [Сотник (підходить до сина, хрестить його і цілує):] *Божє тебе благослови, моя дитино!* (Т. Шевченко); (2) [Ледве (ще тільки)] **благословляло / благословило на світ, безос.** – світало, розвиднялося. *Щебетання ластівок над вікном її невеличкої кімнати будило її [Марію], ще тільки благословляло на світ; а коли з-за обрїю витикалося сонце, воно заставало її вже в полі* (Д. Бедзик); *Ось і сьогодні схпився [Твердохліб], ледве на світ благословило. Дні весняні гарячі, сімба тільки-но почалася, тут не до сну* (І. Цюпа).

Графічні репрезентації комплексних формальних ознак лівої та правої частин, наведені, відповідно, на рис. 3.4 та на рис. 3.5:

Λ(БЛАГОСЛОВЛЯТИ)

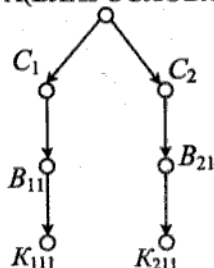


Рис. 3.4. Графічна репрезентація Λ(БЛАГОСЛОВЛЯТИ)

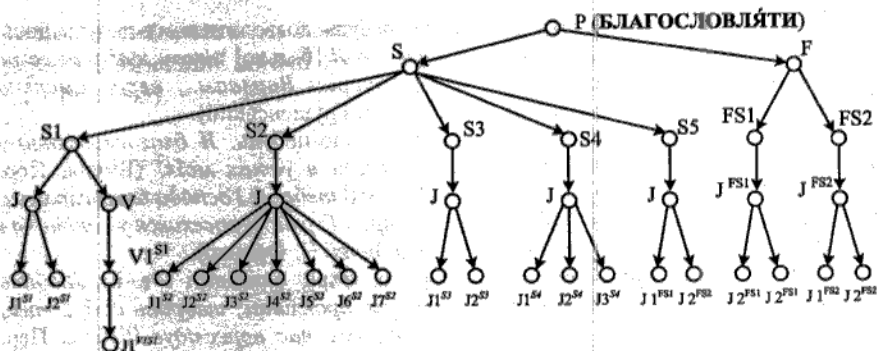


Рис. 3.5. Графічна репрезентація Р(БЛАГОСЛОВЛЯТИ)

Випишемо усі структуротвірні елементи правої частини цієї словникової статті.

1. бібл. Давати милість, благодать, допомогу і т. ін. кому-, чому-небудь; обдаровувати благами (про Бога). *Ти, Господи, благословлятимеш праведного, милістю вкриши його, як щитом!* (Біблія. Пер. І. Огієнка); *Нехай благословляє нас Бог, і всі кінці землі хай бояться Його!* (Біблія. Пер. І. Огієнка); // *кого ким, чим, розм.* Давати, дарувати (про Бога). *Живе Олеся рік і другий .. і четвертий. Благословив Господь діточками: три сини як три соколи* (Марко Вовчок).

2. рел.-церк. Хрестити кого-небудь, проказуючи при цьому молитву, побажання. — *Що ж ти нас не благословляєш, панотче?* — каже Сомко Череваневі (П. Куліш); *Веде Олександра діток, обливаючись сльозами, та все тільки благословляє їх та хрестить...* (Марко Вовчок); *Сиділа [Згода] на балконі нового будинку та все лиш усміхалася та благословила* (І. Франко); *Дай вам Боже.. дочечку до року до вінця благословити...* (Леся Українка); *Благослови ж, матусю, чумакиє На їх непевну і тяжку дорогу* (М. Рильський); *Перед початком роботи делегатів та гостей благословили на добро священники Української автокефальної православної та Української греко-католицької церков* (з газ.); * **Образно.** *Природа .. лагідно благословляла людські веселощі* (Я. Качура).

3. перев. на що. Давати згоду на що-небудь, дозволяти кому-небудь щось. — *Я піду у черницю. — промовляє Катря .. — Я [батько] ніколи тебе на се не благословлю!* (Марко Вовчок); *Вмер старий Варлаам.., перед смертю благословивши свого учня йти в печери* (М. Грушевський).

4. уроч. Прославляти, хвалити кого-, що-небудь. *Я благословлятиму Господа кожного часу, хвала Йому завжди в устах моїх!* (Біблія. Пер. І. Огієнка); *Умийся, серце, щоб пізнала [мати] тебе .. І Господа благословляла за долю добрую твою* (Т. Шевченко); *Старий Гурте?.. поселився у капітана і благословить його ім'я та пильнує його дітей* (І. Франко).

5. Бажати кому-, чому-небудь добра, миру, благодаті і т. ін. *Любіть ворогів своїх, благословляйте тих, хто вас проклинає, творить добро тим, хто ненавидить вас, і моліться за тих, хто вас переслідує* (Біблія. Пер. І. Огієнка); *Не платіть злом за зло або лайкою за лайку, навпаки, благословляйте, знаючи, що на це вас покликано, щоб ви впадкували благословення* (Біблія. Пер. І. Огієнка).

С1: 1. *бібл.* Давати милість, благодать, допомогу і т. ін. кому-, чому-небудь; обдаровувати благами (про Бога). *Ти, Господи, благословлятимеш праведного, милістю вкриєш його, як щитом!* (Біблія. Пер. І. Огієнка); *Нехай благословляє нас Бог, і всі кінці землі хай бояться Його!* (Біблія. Пер. І. Огієнка); // *кого ким, чим, розм.* Давати, дарувати (про Бога). *Живе Олеся рік і другий .. і четвертий. Благословив Господь діточками: три сини як три соколи* (Марко Вовчок).

Л1 ^{С1}: *Ти, Господи, благословлятимеш праведного, милістю вкриєш його, як щитом!* (Біблія. Пер. І. Огієнка).

Л2 ^{С1}: *Нехай благословляє нас Бог, і всі кінці землі хай бояться Його!* (Біблія. Пер. І. Огієнка).

У1 ^{С1}: // *кого ким, чим, розм.* Давати, дарувати (про Бога). *Живе Олеся рік і другий .. і четвертий. Благословив Господь діточками: три сини як три соколи* (Марко Вовчок).

Л1 ^{У1}: *Живе Олеся рік і другий .. і четвертий. Благословив Господь діточками: три сини як три соколи* (Марко Вовчок).

С2: 2. *рел.-церк.* Хрестити кого-небудь, проказуючи при цьому молитву, побажання. – *Що ж ти нас не благословляєш, панотче? – каже Сомко Череваневі* (П. Куліш); *Веде Олександра діток, обливаючись сльозами, та все тільки благословляє їх та хрестить...* (Марко Вовчок); *Сиділа [Згода] на балконі нового будинку та все лиш усміхалася та благословила* (І. Франко); *Дай вам Боже.. дочечку до року до вінця благословити...* (Леся Українка); *Благослови ж, матусю, чумаків На їх непевну і тяжку дорогу* (М. Рильський); *Перед початком роботи делегатів та гостей благословили на добро священники Української автокефальної православної та Української греко-католицької церков (з газ.); ** *Образно.* *Природа .. лагідно благословляла людські веселощі* (Я. Качура).

Л1 ^{С2}: – *Що ж ти нас не благословляєш, панотче? – каже Сомко Череваневі* (П. Куліш).

Л2 ^{С2}: *Веде Олександра діток, обливаючись сльозами, та все тільки благословляє їх та хрестить...* (Марко Вовчок).

Л3 ^{С2}: *Сиділа [Згода] на балконі нового будинку та все лиш усміхалася та благословила* (І. Франко).

Л4 ^{С2}: *Дай вам Боже.. дочечку до року до вінця благословити...* (Леся Українка).

Л5 ^{С2}: *Благослови ж, матусю, чумаків На їх непевну і тяжку дорогу* (М. Рильський).

Ж6⁵²: Перед початком роботи делегатів та гостей благословили на добро священики Української автокефальної православної та Української греко-католицької церков (з газ.).

Л 52: * Образно. Природа .. лагідно благословляла людські веселоці (Я. Качура).

Ж3: 3. перев. на що. Давати згоду на що-небудь, дозволяти кому-небудь щось. – Я піду у черниці. – промовляє Катря .. – Я [батько] ніколи тебе на се не благословлю! (Марко Вовчок); Вмер старий Варлаам.., перед смертю благословивши свого учня йти в печери (М. Грушевський).

Л 53: – Я піду у черниці. – промовляє Катря .. – Я [батько] ніколи тебе на се не благословлю! (Марко Вовчок).

Ж2: 53: Вмер старий Варлаам.., перед смертю благословивши свого учня йти в печери (М. Грушевський).

Ж4: 4. уроч. Прославляти, хвалити кого-, що-небудь. Я благословлятиму Господа кожного часу, хвала Йому завсіди в устах моїх! (Біблія. Пер. І. Огієнка); Умийся, серце, щоб пізнала [мати] тебе .. І Господа благословляла за долю добрую твою (Т. Шевченко); Старий Гурте?.., поселився у капітана і благословить його ім'я та пильнує його дітей (І. Франко).

Л 54: Я благословлятиму Господа кожного часу, хвала Йому завсіди в устах моїх! (Біблія. Пер. І. Огієнка).

Ж2 54: Умийся, серце, щоб пізнала [мати] тебе .. І Господа благословляла за долю добрую твою (Т. Шевченко).

Ж3 54: Старий Гурте?.., поселився у капітана і благословить його ім'я та пильнує його дітей (І. Франко).

Ж5: 5. Бажати кому-, чому-небудь добра, миру, благодаті і т. ін. Любить ворогів своїх, благословляйте тих, хто вас проклинає, творить добро тим, хто ненавидить вас, і молитесь за тих, хто вас переслідує (Біблія. Пер. І. Огієнка); Не платіть злом за зло або лайкою за лайку, навпаки, благословляйте, знаючи, що на це вас покликано, щоб ви впадували благословення (Біблія. Пер. І. Огієнка).

Л 55: Любить ворогів своїх, благословляйте тих, хто вас проклинає, творить добро тим, хто ненавидить вас, і молитесь за тих, хто вас переслідує (Біблія. Пер. І. Огієнка).

Ж2 55: Не платіть злом за зло або лайкою за лайку, навпаки, благословляйте, знаючи, що на це вас покликано, щоб ви впадували благословення (Біблія. Пер. І. Огієнка).

F: ◊ (1) Боже (Господи) благослови – уживається як побажання успіху перед початком якої-небудь справи, схвалення якихось дій. – Благослови сей чесний і важкий хліб у ніч посадити! – Боже благослови! (Г. Квітка-Основ'яненко); [С о т н и к (підходить до сина, хрестить його і цілує):] Боже тебе благослови, моя дитино! (Т. Шевченко); (2) [Ледве (ще тільки)] благословляло / благословило на світ, безос. – світало, розвиднялося. Щebetання ластівок над вікном її невеличкої кімнати будило її [Марію], ще тільки благословляло на світ; а

коли з-за обр'ю витикалося сонце, воно заставало її вже в полі (Д. Бедзик); *Ось і сьогодні схопився [Твердохліб], ледве на світ благословило. Дні весняні гарячі, сімба тільки-но почалася, тут не до сну (І. Цюпа).*

FS1 : ◇ (1) Б́же (Ѓсподи) благослови́ – уживається як побажання успіху перед початком якої-небудь справи, схвалення якихось дій. – Благословіть сей чесний і важкий хліб у ніч посадити! – Боже благослови! (Г. Квітка-Основ'яненко); [С отник (підходить до сина, хрестить його і цілує):] Боже тебе благослови, моя дитино! (Т. Шевченко).

Л^{FS1}: Благословіть сей чесний і важкий хліб у ніч посадити! – Боже благослови! (Г. Квітка-Основ'яненко).

Л^{FS1}: [С отник (підходить до сина, хрестить його і цілує):] Боже тебе благослови, моя дитино! (Т. Шевченко).

FS2 : (2) [Л́две (ще т́льки)] благословля́ло / благослови́ло на світ, безос. – світало, розвиднялося.

Л^{FS2}: Щebetання ластівок над вікном її невеличкої кімнати будило її [Марію], ще тільки благословляло на світ; а коли з-за обр'ю витикалося сонце, воно заставало її вже в полі (Д. Бедзик).

Л^{FS2}: Ось і сьогодні схопився [Твердохліб], ледве на світ благословило. Дні весняні гарячі, сімба тільки-но почалася, тут не до сну (І. Цюпа).

Структурування словникових статей відповідно до сформульованих вище правил надає зручний формальний апарат для аналізу тексту словника та його автоматичного структурування: ідентифікація й виокремлення відповідних структурних елементів і, зрештою, автоматична побудова лексикографічних баз даних (парсинг словника) з метою проведення різнопланових лінгвістичних досліджень і створення нових лексикографічних праць.

3.5. Формалізований підхід до аналізу словникових дефініцій дієслів у СУМі

Теорія лексикографічних систем дозволяє поглибити побудовану структуру правої частини $P(X)$ словникових статей СУМа шляхом застосування процедури рекурсивної редукції до окремих її елементів, зокрема тлумачень. Це досягається у наступний спосіб.

Основною функцією тлумачень є правильне введення кожного слова реєстру в систему мови з урахуванням особливостей його лексичної семантики. Означений процес базується на специфіці внутрішньої структури лексичного значення і зумовлює існування певних норм та способів побудови тлумачення, що регулюються правилами, виробленими лексикографічною традицією.

Моделювання лексичної семантики полягає у виявленні мовних засобів – слів і зв'язків між ними, за допомогою яких можлива декомпресія лексичного значення реєстрових одиниць словника на основі формул

тлумачення, та у встановленні співвідносних за тими чи іншими виокремленими параметрами угруповань.

Формули тлумачення виступають певним чином стандартизованими, відшліфованими структурами, у межах яких функціонування їх компонентів є однозначним і прозорим (принаймні так повинно бути). У структурно-функціональному аспекті формула тлумачення становить автосемантичне речення, що не містить показників синтаксичних зв'язків з іншими реченнями та виступає достатнім для виявлення й однозначної кваліфікації його компонентів.

Якщо в наукових дослідженнях [165; 243], пов'язаних, як правило, з ідеографічним вивченням лексики, компонентом лексичного значення (семою, семантичною ознакою, маркером, елементарним смислом) прийнято вважати окреме слово в кожній самотійній рубриці словникової статті, окрім «нульових компонентів, чи зв'язок» (який стосується, який має, який характеризується, який-небудь, той, щось, предмет, властивість, стан, відношення та ін.), то у нашому дослідженні із застосуванням прийомів компонентного аналізу семантичної структури речення виділилися логічні блоки дефініцій. Слід зауважити, що компонентний аналіз у класичному вигляді представляє собою саме виділення сутнісних характеристик об'єктів реальної дійсності, засобом номінації яких служить те чи інше слово.

Наприклад, у третьому значенні дієслова **ІТИ (ЙТИ)** «Виходити, вирушати в певному напрямі до кого-небудь» наявні такі компоненти: семантичний предикат *виходити* та *вирушати*; семантичний конкретизатор *місця в певному напрямі*; семантичний об'єкт *до кого-небудь*.

У тлумаченні другого відтінку п'ятого значення дієслова **СВІТІТИСЯ** «Виявляти своїм зовнішнім виглядом задоволення, радість, щастя (про людину)» – семантичний предикат *виявляти*; семантичний конкретизатор *способу своїм зовнішнім виглядом*; семантичний конкретизатор об'єкта *задоволення, радість, щастя*; семантичний конкретизатор суб'єкта *людина*.

У тлумаченні першого значення дієслова **КАТАТИ** «Возити когось на чому-небудь або в чомусь для розваги, на прогулянку» – семантичний предикат *возити*, семантичний конкретизатор об'єкта *когось*, семантичний конкретизатор засобу *на чому-небудь або в чому-небудь*, семантичний конкретизатор мети *для розваги, на прогулянку*.

Як свідчить аналіз досліджуваного матеріалу, формула тлумачення дієслівних лексем складається з компонентів, які в її межах різняться за своїм функціональним навантаженням та характером взаємовідношень між собою. Такі компоненти об'єднуються у два блоки: ідентифікатор (семантична тема [295:10]) та конкретизатори (диференційні семи). Інші терміни на позначення компонентів тлумачення: парадигматичний ідентифікатор та диференційний елемент [175]; основна та додаткова частини

[170:12]; основні відповідники (вільні і зв'язані) та лексикографічні конкретизатори (модифікатори і конкретизатори) [217:131; 176–178].

Ідентифікатором у дієслівних дефініціях виступає семантичний предикат – інше дієслово, за допомогою якого тлумачиться реєстрове. Це дієслово ми будемо називати терміном, запропонованим Д. М. Шмельовим, семантична тема (інші найменування такі: архілексема, класема (Е. Косеріу), тематичний доміант (В. Морковкін), опорне слово, ключове слово та ін. У межах конкретно взятого значення семантична тема як ідентифікатор забезпечує об'єднання навколо себе інших компонентів, що становлять диференційний блок тлумачення.

Семантична тема, будучи стрижневою в семній структурі лексичного значення й одночасно забезпечуючи входження кожного лексико-семантичного варіанта (ЛСВ) у систему взаємовідношень з ЛСВ інших слів, як правило, конкретизується диференційними семами, регулярними за своїм проявом, серед яких є більш чи менш регулярні. Слід зауважити, що інколи визначення тієї чи іншої диференційної семи не є однозначним.

Отже, зважаючи на особливості дієслівної семантики, останню можна представити у вигляді ідеографічно-таксономічного формату запису дієслівних формул тлумачення. Відзначено констатується такими твердженнями:

1) речення, якими представляються повноцінні (невідсильні) формули тлумачення, є однозначними репрезентантами семантичних станів лексеми;

2) у структурі речень, якими представляються повноцінні (невідсильні) формули тлумачення, міститься усе необхідне для характеристики даного семантичного стану.

Визнання цих двох тверджень зумовлює можливість декомпозиції формули тлумачення дієслів на окремі компоненти – детермінанти значення, характеристичні та визначення яких здійснимо шляхом інтерпретації множини формул тлумачення як своєї лексикографічної системи і застосування до неї процесу рекурсивної редукції. Зазначений процес проведемо, спираючись на такі положення:

1) правило декомпозиції формули тлумачення на компоненти є одним для усіх дієслів;

2) кожен компонент має свій узагальнений елементарний смисл і у певному смислі є семою;

3) серед усіх компонентів формули тлумачення виділяється головний, який є обов'язковим і складає своєрідний центр семантичного стану – це семантична тема (*ST*). Для кожного семантичного стану існує одна і тільки одна *ST*;

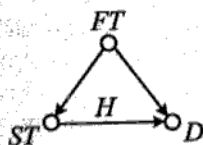
4) усі інші компоненти детерміновані семантичною темою і можуть бути факультативними – це диференційні семи; множину диференційних сем позначимо літерою *D*;

5) диференційні семи перебувають у взаємодії із семантичною темою та між собою у такий спосіб, що кожна з них зменшує обсяг семантичного поля, визначеного семантичною темою;

6) набір компонентів дієслівних формул тлумачення становить універсальну лінійну структуру;

7) компоненти формули тлумачення є функціональними, проте виділяються уніфіковані форми для вираження того чи іншого компонента; можливе перефразовування компонентів формули тлумачення у напрямку оптимізації та уніфікації.

У такий спосіб множина формул тлумачення дієслів *FT*, розглядувана як лексикографічна система, набуває розвинення:



(3.34)

де множина семантичних тем *ST* відіграє роль реєстрової (формальної) частини лексикографічної системи *FT*, а множина диференційних сем *D* – роль її інтерпретаційної (змістової) частини; *H*, як завжди, здійснює зв'язок між ними.

Формула (3.34) набуває конкретного змісту при підстановці до неї конкретної дієслівної лексеми *X*:

$$FT(X) = ST(X) \xrightarrow{H} D(X) = ST(X) + D(X). \quad (3.35)$$

Шляхом аналізу словникових дефініцій дієслова у СУМі виділено типи диференційних сем, характерні для дієслівних формул тлумачення та встановлено типові структури усіх компонентів, включаючи *ST*. Виявлено та ідентифіковано набір із 10 диференційних сем: *Dsbj* – суб'єкт; *Dobj* – об'єкт; *Dadr* – адресат; *Dinstr* – спосіб, засіб, матеріал; *Dloc* – середовище, місце; *Dtemp* – час; *Daim* – мета; *Dcond* – умова; *Dcaus* – причина; *Dmsr* – міра, інтенсивність. Зазначений набір є інваріантним, унаслідок чого будь-яка дієслівна формула тлумачення розкладається у лінійний (табличний) формат зі структурою, представленою у табл. 3.16:

Таблиця 3.16. Формат дефініцій дієслова

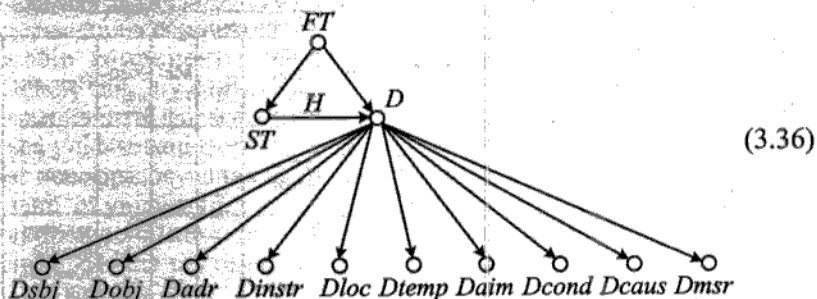
Семантична тема	Диференційні семи									
	Суб'єкт	Об'єкт	Адресат	Спосіб, засіб	Середовище, місце	Час	Мета	Умова	Причина	Міра, інтенсивність
1	2	3	4	5	6	7	8	9	10	11

Як приклад наведемо запис у табличному форматі формул тлумачення дієслова ЛЕТИТИ (табл. 3.17):

Таблиця 3.17. Формат дефініцій дієслова ЛЕТІТИ

№ семантичного стану	Семантична тема	Диференційні семи										
		Суб'єкт	Об'єкт	Адресат	Спосіб, засіб	Середовище, місце	Час	Мета	Умова	Причина	Міра, інтенсивність	
1.	пересуватися	(про птахів, комах і т. ін.)			за допомогою крил	в повітрі						
1. //1.	пересуватися					в повітрі				завдяки певній зовнішній силі (вітру, поштовху і т. ін.)		
1. //2.	пересуватися, переміщатися	(про літаки, ракети і т. ін.)				в повітрі						
1. //3.	пересуватися	(про пасажирів або пілота)			на літальних апаратах	в певному напрямі						
1. //4. перен.	линути, розноситися	(про слова, звуки тощо)				в просторі						
1. //5. перен.	ширитися, передаватися, пересилатися				певним способом (поштою, радіо тощо)							
2.	те саме, що ПАДАТИ 1, 2											
3. перен.	переміщатися; мчати; нестися (у 1 знач.)					(по землі, воді і т. ін.)						дуже швидко
4. перен.	проходити	(про час)			непомітно							швидко
5. перен.	линути		до когось, чогось, чого-небудь		думками, почуттями, помыслами, душею							

У такий спосіб формула (3.34) набуває такого розвинуття:



Отже, структура узагальненої формули тлумачення для дієслова представляється у такий спосіб:

$$FT = ST + D = ST + Dsbj + Dobj + Dadr + Dinstr + Dloc + Dtemp + Daim + Dcond + Dcaus + Dmsr. \quad (3.37)$$

Порядок слідування компонентів тут не відображено (так само, як і синтаксичну структуру формули тлумачення).

Запропонований формат дієслівних формул тлумачення виступає простішим об'єктом, ніж словникова дефініція, що робить його доступним для автоматизованого дослідження.

У працях [256; 225] здійснено лінгвістичний аналіз одержаного формату та виявлено типові реалізації його окремих компонентів — диференційних сем у лексикографічній системі СУМ-11. Проведена систематизація дозволила здійснити лексико-семантичну стратифікацію дієслівної лексики. Зокрема показано, що найбільш абстрактними у значенні дієслова є несубстанційні семи, що визначають лексичне значення, властиве саме дієслову як окремому лексико-граматичному класу. Ними виступають семи *дія*, *стан*, *відношення*. На нижчому рівні ієрархії класи дієслів *дії*, *стану* та *відношення* за специфічними семами, репрезентованими в структурі формули тлумачення семантичною темою, конкретизуються щодо характеру дії, стану та відношення і поділяються на підкласи — лексико-семантичні групи (дієслова руху, мислення, мовлення та ін.). Безпосереднє звернення до аналізу семантичних тем показує, що об'єднання аналізованих дієслів у відповідні лексико-семантичні групи забезпечується найчастотнішими серед них, які співвідносяться з поняттями руху, мислення, мовлення та ін. відповідно. Наприклад, інтегральною для окремих лексико-семантичних

варіантів дієслів *бігати, відходити, іти (йти), лазити, літати, летіти, плавати, плисти (пливти, плунути), повзати, проїжджати, стрибати, ходити* та под. виступає *ST пересуватися*; для окремих лексико-семантичних варіантів дієслів *агакати, базікати, брехати, верзти, вимовляти, гомоніти, горлати, додавати, защебетати, лепетати, підказувати, повторювати, промовляти, твердити* та под. — *ST говорити*.

Метод форматного представлення словникових дефініцій служить надійним апаратом для встановлення співвідносних множин лексики і за диференційними семами. Останні, властиві окремим ЛСВ, як правило, не підпорядковуються певній визначеній ієрархії, а перебувають між собою у відношенні рівноправності як лінійні елементи формату.

Переведення словникових тлумачень в ідеографічно-таксономічний формат уможливило виявити скінченні списки універсальних семантичних ознак різної природи, що співвідносяться з певним атрибутом формату (семантичною темою чи диференційними семами *суб'єкт, об'єкт, адресат, спосіб, засіб, середовище, місце, час, мета, причина, умова, міра, інтенсивність*) і служать елементами метамови лексикографічного опису лексичної семантики усіх ЛСВ дієслова.

Саме виділені атрибути формату дозволяють проведення багатопараметричних лексико-граматичних класифікацій дієслова. В цілому ці класифікації можна об'єднати у три групи:

- 1) за атрибутами (одним чи кількома) ідеографічно-таксономічного формату;
- 2) за конкретними значеннями, які може приймати атрибут (один чи кілька) ідеографічно-таксономічного формату;
- 3) за атрибутом (одним чи кількома) та значеннями іншого (чи інших) атрибута ідеографічно-таксономічного формату.

Аналіз словникових тлумачень, переведених у табличний формат, свідчить, що в основі лексико-семантичної системи сучасної української мови лежить певна кількість універсальних семантичних ознак, поєднання яких між собою утворюють масив реальних лексичних значень дієслів.

Важливо, що використання введеного формату при укладанні дієслівних словникових статей дозволяє уніфікувати (там, де це можливо й доцільно) тлумачення окремих значень, а також сприяє якомога повнішому описові семантики тих чи інших лексем, увиразнюючи семантичну структуру дієслів.

3.6. Формалізований підхід до аналізу словникових дефініцій у СУМі

Дослідження формул тлумачення іменників здійснювалося у межах поділу ресурсного складу СУМ-11 на лексико-семантичні групи (ЛСГ). Для прикладу подасмо аналіз структури формул тлумачення для двох іменникових семантичних груп: назв знарядь праці та споруд.

Такі характерні ознаки ЛСГ, як категоріальна спільність лексичних одиниць, у межах якої встановлюється семантичний зв'язок між словами та їхніми значеннями, ієрархічний порядок їхньої організації, відносна автономність одиниць, що входять до її складу, безперервність смислового простору [190:122], дозволяють провести системний лексико-семантичний аналіз іменників на позначення назв знарядь праці та споруд. Семантичні особливості цих лексем виявляються у предметному (денотативному) значенні, формування якого здійснюється на основі переліку ключових ознак денотата. Тому семантика досліджуваних лексем орієнтується на об'єктивно виділений фрагмент зовнішнього світу, що дозволяє здійснити прозору та однозначну ідентифікацію семантичних компонентів у структурі їх значення.

Застосування теорії Л-систем при дослідженні тлумачень іменникових лексем на позначення назв знарядь праці та споруд дозволило виявити системно-структурні основи функціонування їх лексичної семантики у лексикографічному представленні в СУМ-11.

Права частина словникової статті іменників $P(X)$, зокрема тлумачення, що міститься в ній, розглядається з точки зору набору структуротвірних елементів, які характеризуються узагальненим змістом та логіко-синтаксичними зв'язками, що встановлюються між ними. Це дозволяє формалізувати структуру тлумачень іменників досліджуваних ЛСГ, вивести лексикографічні формули їх тлумачень та створити модель лексикографічної системи, що відображає феномен лексичної семантики таких слів.

Вихідним етапом дослідження тлумачень як способу опису лексичної семантики мовних одиниць у СУМ-11 є вивчення їх структуротвірних елементів, які впорядковуються на рівні логічних відношень і функціонують у кореляції з семантичними компонентами структури значення.

При дослідженні логічної структури тлумачень було встановлено характер зв'язків між елементами словникової дефініції [133:143]. Ними визначено родо-видові та партитивні зв'язки, що становлять логіко-поняттєву основу тлумачень.

На рівні логічних відношень лексикографічне тлумачення формується за допомогою обов'язкових структурних елементів, – родової та видової ознак, поєднаних зв'язком певного змісту. У структурі тлумачення такі елементи виконують класифікаційну функцію: перший, відповідно, вказує на поняттєвий клас, до якого належить описуване поняття реєстрового слова, другий – визначає і обумовлює родову ознаку. Загальна структура тлумачення на логічному рівні уточнюється складовими елементами та зв'язками, що встановлюються між ними, і реалізується у вигляді:

$$T(x) = Dfd \cap Rn Dfn (K_1 + K_2), \quad (3.38)$$

де $T(x)$ – тлумачення слова x , Dfd (дефінієндум) – тлумачне слово, Dfn (дефінієнс) – опис поняття тлумачного слова, $K_1 + K_2$ – складові структурні компоненти Dfn – відповідно родова та видова ознаки, $\cap Rn$ – логічний зв'язок, яким поєднуються Dfd та Dfn .

Структурні компоненти, виділені в тлумаченнях на рівні логічних відношень, закономірно функціонують як певні елементи, що характеризуються типовою змістово-функціональною роллю: родова ознака вказує чи визначає поняттєвий клас, та видова ознака виступає уточнювальною ознакою щодо родової. Ці елементи є структуротвірними і обов'язковими у формуванні тлумачення іменникових лексем на позначення знарядь праці та споруд.

При подальшому аналізі тлумачень іменникових лексем досліджуваних ЛСГ було встановлено, що виділені структурні елементи тлумачення на рівні логічних відношень із вказівкою на родову та видову ознаки співвідносяться із семантичними компонентами – відповідно, семантичною темою та компонентом функціонального призначення, які на рівні лексичного значення є обов'язковими і центральними семантичними компонентами тлумачень іменників на позначення знарядь праці та споруд.

Семантична тема як ідентифікуючий елемент являє собою узагальнююче поняття, під яке підводиться видове поняття або поняття зі значенням частини цілого. При цьому семантична тема на рівні логічної організації тлумачення співвідноситься з родовою ознакою як його обов'язковим елементом. З точки зору поняттєвої організації тлумачення цей семантичний компонент можна кваліфікувати обов'язковим структуротвірним, без якого визначення як логічна категорія не може бути реалізоване.

У представлених в СУМ-11 тлумаченнях іменникових лексем ЛСГ на позначення назв знарядь праці та споруд розрізняються семантичні теми різних рівнів ієрархії, що стало наслідком появи нових слів на означення реалій праці, які з'явилися впродовж розвитку та удосконалення виробничої діяльності людини.

На основі аналізу семантичної структури тлумачень іменникових лексем на позначення знарядь праці у цій групі лексем виділено такі семантичні теми чотирьох рівнів узагальнення:

1-й рівень: поняття найвищого рівня узагальнення – знаряддя, яке охоплює весь даний клас лексем;

2-й рівень: за сутнісною ознакою знарядь праці виділяються семантичні теми нижчого рівня – *машина, пристосування, апарат, механізм, інструмент, прилад, пристрій*. На їх основі у лексико-семантичній групі на позначення назв знарядь праці виділяються відповідні мікрогрупи. Так, родові поняття *пристрій* та *прилад* об'єднують найбільшу кількість лексем – 260, меншим складом характеризуються мікрогрупи, утворені на основі таких семантичних тем, як *машина* – 164, *інструмент* – 83, *апарат* – 68, *пристосування* – 34 та *механізм* – 23. На основі спільного родового ідентифікатора між виділюваними в такий спосіб лексемами у межах мікрогруп встановлюються синонімічні зв'язки;

3-й рівень: семантичні теми нижчого рівня узагальнення, які на основі методу ступінчастої ідентифікації [181:84–85] співвідносяться з поняттями вищого рівня: *лом, лампа, струнотримач, молоток, ніж, кран, важіль, ліхтар, голка, засувка, циркуль, снаряд* та ін. Наприклад: *кронциркуль* – маленький циркуль для креслення дуг радіусом до 1 мм; *колий* – ніж, яким колють свиней; *квянка*² – дерев'яний молоток, яким користуються в столярстві; *трубоукладач* – пересувний підйомний кран на гусеничному тракторі для укладання труб у траншеї;

4-й рівень: семантичні теми найнижчого рівня узагальнення. Сюди належать ідентифікатори типу: *брус, трубочка, балка, дошка, жердка, планки, тичка, ланцюг, трос, паличка, стрижень, держак, ручка, розвилка* та ін. Наприклад: *кітві* – металевий стрижень з лапами, що зачіплюється за ґрунт і утримує на місці судно, пливучий маяк і т. ін.; *якір*; *рогátка* – дерев'яна розвилка з прив'язаною до обох кінців гумовою стьожкою для метання чого-небудь; *вішал* – закріплена на стовпах або підвішена на мотузках чи дротинах тичка для сушіння чого-небудь; *жердка*.

У лексико-семантичній групі на позначення назв споруд виділяються ідентифікатори трьох рівнів узагальнення:

1-й рівень представлено узагальнюючим поняттям *споруда*, під яке підводиться весь клас даних слів;

2-й рівень характеризується наявністю двох родових ідентифікаторів: *будівля* та *будинок*. Так, ідентифікатором *будівля* позначається значення лексем: *хлів, альтанка, стодола, балаган, клуня, дім, будинок*, тоді як по-

няття *будинок* також виявляється родовим і ним визначаються лексеми: *садиба, будка, готель, театр, кімната*;

3-й рівень представлено семою-ідентифікатором нижчого рівня узагальнення *приміщення*, який, у свою чергу, визначається поняттям вищого рівня узагальнення *будівля*. За наявності ідентифікатора *приміщення* виділяється синонімічна підгрупа лексем: *помешкання, майстерня, повітка, гуртожиток, крамниця, кабіна, склеп, школа, цех* та ін.

Таким чином, семантична тема як ідентифікуючий елемент у тлумаченнях досліджуваних іменників кваліфікується обов'язковим семантичним компонентом, який забезпечує номінацію вищого порядку. Він ідентифікує, вказує на клас предметів, до яких належить предмет, позначений тлумачним словом.

У більшості досліджуваних слів тлумачення формуються на основі однієї семантичної теми. В окремих тлумаченнях можна спостерігати функціонування кількох семантичних тем одночасно. Як переконують результати дослідження, такі подвійні конструкції можна вважати єдиним багаточленним компонентом лише за умови їх змістової подібності чи семантичної еквівалентності. У таких випадках маємо справу з дво-, тричленною структурою компонента семантичної теми, утвореного набором синонімічно споріднених ідентифікаторів із вказівкою на родові характеристики: *стабілізатор – пристрій, прилад, що надає стійкої рівноваги тілу, яке рухається; очісник – пристрій, механізм для очищення чого-небудь; розширник – пристрій, інструмент для розширення чого-небудь*. Реалізація компонента «семантична тема» кількома ідентифікаторами є свідченням того, що мікрогрупи, утворювані на основі таких ідентифікаторів у системі мови, перетинаючись, накладаються одна на одну. Зонами перетину окремих мікрогруп є їх периферійні області.

Автоматизовані вибірки за компонентом «семантична тема» дозволили виявити неточності та деякі некоректності у формулюванні окремих тлумачень у 11-томному тлумачному Словнику української мови. Йдеться, зокрема, про поєднання у структурі семантичної теми ідентифікаторів різних семантично віддалених груп. Порівняйте: *дім – приміщення, люди, що в ньому живуть, та їхнє господарство; сповільнювач – пристрій або речовина, що сповільнює рух, дію чого-небудь; загородження – спеціальна перешкода, певна споруда або засоби руйнування для перекриття підступів до чого-небудь; фарботёрка – машина, дощка або чашка для розтирання фарб; уповільнювач – пристрій або речовина для уповільнення дії чого-небудь; міновоз – машина або віз для перевезення мін або інших боєприпасів, а також той, хто їх возить, зерносушарка – установ-*

ка, механічний пристрій або спеціально обладнане приміщення для сушіння зерна (у 1, 2 знач.) з доведенням його до кондиційної вологості; рулетка – спеціальний пристрій для певного роду азартної гри, а також сама ця гра; загородження – спеціальна перешкода, певна споруда або засоби руйнування для перекриття підступів до чого-небудь; запарник – казан або якийсь пристрій, признач. для запарювання чого-небудь; фільтр – пристрій, прилад або речовина для очищення рідини, газу тощо від непотрібних домішок; салон – приватний дім, у якому постійно збираються гості – особи аристократичного чи буржуазного кола; саме світське товариство, що збирається в такому домі.

Несумісність таких ідентифікаторів у тлумаченні одного поняття вказує на необхідність розмежування їх за тлумаченнями окремих лексико-семантичних варіантів з відповідними родовими семами.

Семантичний компонент функціонального призначення (умовн. позн. – *F*) на рівні логічних відношень співвідноситься з елементом видової ознаки.

Відповідно до функціонування цього компонента на рівні логічних відношень як видової ознаки, яка диференціює родову ознаку, компонент функціонального призначення кваліфікуємо диференційним компонентом, що конкретизує семантичну тему стосовно ознаки різноманітного призначення, застосування.

Семантичний компонент функціонального призначення реалізується набором конструкцій, утворених на основі поєднання відносних займенників / прийменників з дієсловом чи прийменника з віддієслівним субстантивом, який безпосередньо вказує на дію, конкретизує процес, із яким пов'язується функціонування предмета. Це такі конструкції:

«прийм. для / (признач.) для + віддієслівний іменник»: бурякокопач – знаряддя для викопування буряків; цибулесховище – відповідно обладнане приміщення для зберігання цибулі; клуня – будівля для зберігання снопів, сіна, полови тощо, а також для молотьби, віання й т. ін.; стодола; дергач – знаряддя для витягування цвяхів; волокуша – знаряддя для згортання й транспортування сіна та соломи; мотівіло – знаряддя, пристосоване для змотування пряжі, шовку тощо; гучномовець – пристрій для відтворення й підсилення звуку радіопереда; заглушка – пристрій, признач. для закривання отвору в будь-якому приладі, апараті, машині; зачіпка – пристрій, признач. для зачеплення чого-небудь; гармата – артилерійська довгоствольна вогнепальна зброя з положою траєкторією, признач. для ураження живої сили і бойової техніки противника, руйнування оборонних споруд тощо.

Така конструкція є найпоширенішою для тлумачень іменникових лексем, семантика яких визначається функціональним призначенням, зокрема для лексем семантичних груп на позначення назв знарядь праці та споруд.

Також зустрічаються й такі описові, рідше вживані способи зв'язку компонента функціонального призначення з семантичною темою:

– «віднозн. займ. який / що (використовується) + прийм. для + віддієсл. іменник»: *зміншувач – прилад, який використовується для зменшування зображення або дає зменшене зображення чого-небудь; шило – інструмент у вигляді загостреного металевого або кістяного стержня з ручкою, який використовується перев. для проколювання дірок; шприц – хірургічний інструмент у вигляді циліндра з порином та порожнистою голкою, який використовується для впорскування в організм лікарських речовин, а також для відсмоктування звідти рідини;*

– «віднозн. займ. який / що (застосовується) + прийм. для + віддієсл. іменник»: *альтиметр – прилад, що застосовується для визначення висоти місцевості або висоти польоту; висотомір; лійка¹ – такий же конус з трубкою, що застосовується для засипання в що-небудь сипучих речовин; фумель – шевський інструмент, який застосовується для згладжування та вирівнювання поверхні.*

Вибірki за цим компонентом указують на наявність найбільш частотних і характерних для кожної з аналізованих ЛСГ видів призначення. Так, наприклад, для іменникового класу лексем на позначення знарядь праці найхарактернішим стає призначення, що відображається таким значенням дієслів: *вимірювати* (124 лексеми), *тримати* (у знач. підтримувати) (38 лексем), *подавати* (37 лексем), *різати* та похідні (36 лексем), *визначати* (33 лексеми), *очищувати* (33 лексеми), *з'єднувати*, *скріплювати* (28 лексем), *перетворювати* (29 лексем), *записувати* (24 лексеми), *обробляти* (23 лексеми), *ловити* (23 лексеми), *робити*, *виробляти* (20 лексем), *передавати* (17 лексем), *піднімати* (16 лексем).

Для лексем на позначення споруд найхарактернішим стає компонент функціонального призначення, що реалізується за допомогою дієслів: *зберігати* (70 лексем), *готувати*, *виготовляти* (23 лексеми), *тримати* (у знач. утримувати) (19 лексем), *захищати* (18 лексем) *проживати* (7 лексем), *відпочивати* (6 лексем), *складати* (6 лексем) та ін.

У результаті декомпозиції тлумачень досліджуваних лексем встановлено, що семантична структура досліджуваних лексем дворівнева: поряд з компонентами вищого рівня, – семантичною темою і компонентом функціонального призначення виділяються компоненти нижчого рівня узагальнення, які функціонують у ролі уточнювальних компонентів при центральних.

У свою чергу, компоненти нижчого рівня узагальнення розмежовуються на такі, що формуються при обов'язковому компоненті – семантичній темі та є окремим набором при компоненті функціонального призначення – як елементи з поширюючим значенням.

При семантичній темі як уточнювальні виділяються атрибутивний компонент і компонент на позначення місця.

Атрибутивний компонент (умовн. позн. – *atr*) виступає у функції означення, набуваючи властивості доповнювати семантику родового поняття, уточнювати його за якоюсь ознакою. Наприклад: *газонокосарка – невеличка ручна косарка, якою скошують траву на газонах; піддирáč – інструмент у вигляді клинка для видобування ламких порід; мантáчка – вузький дерев'яний плоский брусок для гостріння коси, укритий шаром смоли з піском.*

Цей компонент виявляє класифікаційні властивості, за якими серед досліджуваних груп лексем можуть бути виділені іменникові групи за тим чи іншим критерієм атрибутивності: за формою, висотою, шириною, об'ємом та ін.

Компонент на позначення місця (умовн. позн. – *locat*) також є уточнювальним компонентом семантичної теми, він виконує функцію ознаки з вказівкою (лише в одному зі своїх значень – як вказівка) на місце розташування предмета, де знаходиться описуваний предмет. Наприклад: *зерновловлювач – пристрій на збиральних машинах для уловлювання зерна, насіння, що обсіпається під час збирання врожаю; трубоукладáč – пересувний підйомний кран на гусеничному тракторі для укладання труб у траншеї; гідрáнт – водорозбірний пристрій на водопровідній мережі для подачі води, здебільшого при гасінні пожеж; трюм – внутрішнє приміщення корабля між палубою і днищем, де встановлюють машини, механізми, поміщають вантажі і т. ін.; клас – кімната в школі, гімназії, де здійснюється навчальний процес.*

Уточнювальними компонентами, які стосуються компонента функціонального призначення і об'єднуються навколо нього, є семантичні компоненти на позначення місця, сфери, часу, умови, мети, інструмента, об'єкта і суб'єкта, засобу, способу і напрямку.

На семантико-синтаксичному рівні тлумачення функціонування таких семантичних компонентів зумовлюється валентісно-інтенційними властивостями дієслова, яким реалізується компонент функціонального призначення. Такий семантичний зв'язок між компонентом функціонального призначення та уточнювальними компонентами у тлумаченнях досліджуваних лексем зображено у табл. 3.18:

Таблиця 3.18. Відображення семантичного зв'язку між компонентом функціонального призначення та деякими диференційними уточнювальними компонентами у тлумаченні досліджуваних лексем

Лексема	Компонент семантична тема (ST)	Компонент функц. призн. (F)	Пропозиції дієслова в конструкції компонента функціонального призначення					
			Об'єкт (obj)	Мета (aim)	Час (time)	Умова (cond)	Місце (locat.)	Спосіб (mod)
Віварій	приміщення	для утримування	тварин	із навчальною та експериментальною метою			в умовах, близьких до природних	
Кліть	пристрій	для підмання й спускання	людей, вагонеток із копалиною, породою та різними матеріалами				по шахтовому стовбуру	
Маточник	пристосування	яким доглять	бджолину матку		під час роїння			
Реактор	апарат	для проведення	хімічних реакцій				при певних температурах і тисках	
Фотостат	автоматичний апарат	для копіювання	документів, малюнків, текстів книжок і т. ін.					фотографічним способом

Як видно з табл. 3.18, дієслово або віддієслівний іменник у конструкції вираження компонента функціонального призначення виявляють властивість розгортати певну кількість актантних позицій обов'язкового характеру, що залежать від його семантики. Незважаючи на повний пропозитивний набір актантів дієслова чи віддієслівного іменника, їх добір у тлумаченні вмотивується лінгвістичною компетенцією лексикографа.

Формування і функціонування диференційних компонентів у структурі тлумачення у тісному зв'язку з семантичною темою та компонентом функціонального призначення дозволило розглядати тлумачення як таке, що формується на основі двох семантичних комплексів:

- 1) родової ознаки з семантичною темою та поширюючими її компонентами – атрибутивним та на позначення місця (у відповідному значенні і реалізації);
- 2) видової ознаки з компонентом функціонального призначення та його уточнювальними диференційними компонентами: на позначення

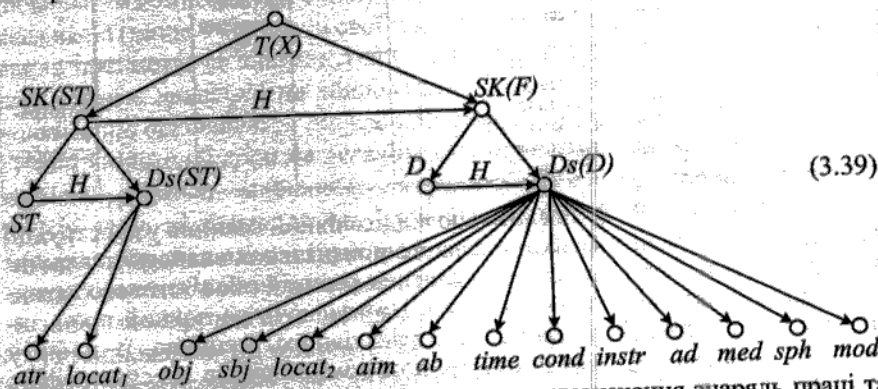
місця (у відповідному значенні й реалізації), об'єкта, суб'єкта, інструмента, засобу, часу, умови, способу, мети, сфери, напряму дії.

Отже, структуру тлумачення на лексико-семантичному рівні можна відобразити за допомогою лінійного формату, представленого у табл. 3.19:

Таблиця 3.19. Лінійний формат структури значення іменників на позначення зрядь праці та споруд

		Іменникове тлумачення													
Ресстр. слово	Семантичний комплекс	Семантичний комплекс													
Центральні компоненти	Компонент семантична тема (ST)	Компонент функціонального призначення (F)													
Дифер. компоненти		atr	locat ₁	obj	sub _j	aim	instr	med	ab	ad	sph	locat ₂	time	cond	mod

Відповідно до узагальненого змісту семантичних компонентів та зв'язків, що встановлюються між ними у структурі тлумачення досліджуваних іменників, виведено модель лексикографічної системи, що відображає феномен лексичної семантики таких одиниць:



де $T(X)$ – тлумачення іменникової лексики на позначення зрядь праці та споруд; $SK(ST)$ – семантичний комплекс із центральним компонентом семантична тема; $SK(F)$ – семантичний комплекс із центральним інтегральним компонентом функціонального призначення; ST – компонент диференційним компонентом семантична тема; $Ds(ST)$ – уточнювальні диференційні компоненти семантичної теми: атрибутивний компонент (atr) та компонент на позначення місця при семантичній темі ($locat_1$); D – дієслово або віддієслівний субстантив, на основі якого реалізується компонент функціонального призначення;

$Ds(D)$ — уточнювальні диференційні семантичні компоненти стосовно дієслова: компоненти на позначення об'єкта (*obj*), суб'єкта (*subj*), місця (*locat₂*), мети (*aim*), напрямку: звідки (*ab*), куди (*ad*), часу (*time*), умови (*cond*), інструмента (*instr*), засобу (*med*), сфери (*sph*), способу (*mod*); H — оператор, що відображає зв'язок між семантичними комплексами $SK(ST)$ та $SK(F)$.

У виведеній моделі структури лексичної семантики іменникових лексем на позначення знарядь праці та споруд функціонування інформаційних елементів, за допомогою яких здійснюється відображення лексичної семантики у лексикографічному описі, розгортається за принципом рекурсивної редукції, який є наскрізним у відображенні лінгвістичної інформації у структурі СУМ-11. Зокрема творення тлумачень досліджуваних лексем моделюється за принципом рекурсивної редукції другого порядку, що виявляється у двоступеневому поширенні семантичних явищ предмета дослідження. Перший рівень представлено виділенням центральних компонентів, які формують семантичні комплекси у структурі тлумачення досліджуваних лексем. При цьому між ними встановлюється зв'язок, що набуває атрибутивного змісту: атрибутом семантичної теми виступає компонент функціонального призначення. На другому рівні даної моделі виділяємо диференційні компоненти, що функціонують відносно центральних компонентів як їх семантичні поширювачі.

Виходячи з виявленого набору семантичних компонентів значення іменникових лексем на позначення знарядь праці та споруд, було виведено формулу, на основі якої реєстрові одиниці досліджуваних лексико-семантичних груп реалізують своє значення. Структуру значення іменників лексико-семантичної групи на позначення знарядь праці можна представити за допомогою формули:

$$T(x)=[atr + locat_1]+ ST + F + [obj + subj + instr + med + mod + locat_2 + sph + ab + ad + aim + cond + time]. \quad (3.40)$$

Для іменників лексико-семантичної групи на позначення споруд така формула набуває спрощення⁸:

$$T(x)=[atr + locat_1] + ST + F + [obj + mod + locat_2 + ab + ad]. \quad (3.41)$$

Квадратні дужки в обох формулах вказують на можливість набуття семою факультативного значення.

Представлені формули свідчать про те, що структура значення іменникових лексем на позначення знарядь праці та споруд характеризується різним набором семантичних компонентів, зумовлюючись відмінностями денотативного значення лексем відповідних лексико-семантичних груп.

⁸ У формулі тлумачення таких лексем фіксуємо лише характерні семантичні компоненти значення, не враховуючи нетипові рідковживані компоненти.

ГОЛОВНІ ЗАСАДИ УКЛАДАННЯ СУМ-20

4.1. Основні принципи лексикографування

При укладанні основної частини СУМ-20 застосовано такі лексикографічні принципи:

1. *Максимально повне збереження і використання теоретичних та практичних надбань вітчизняної лексикографії, втілених у структурі та корпусі СУМ-11.* Зокрема, матеріали зі вступної статті до СУМ-11 використано і в цьому виданні. Крім того, на основі теорії лексикографічних систем проведено велику роботу з вивчення імпліцитної структури СУМ-11, що дало можливість установити ряд нових лінгвістичних фактів, викладених у попередньому розділі. Повністю використано реєстр СУМ-11. Кількість вилучень із реєстру становить лише близько 800 одиниць із понад 134 тисяч. Це такі лексеми, що вийшли з сучасного вжитку (до того ж немає підстав сподіватися, що вони активізуються у певній осяжній перспективі); росіянізми, замінені на питомі українські слова та слова, написання яких змінилося. Значною мірою збережено і семантичну інтерпретацію лексем, особливо для службових частин мови. При цьому виправлені помилки та неточності, неминуче наявні у лексиконах такого великого обсягу.

2. *Звільнення СУМа від рудиментів тоталітарного режиму; деідеологізація лексикографічного матеріалу; забезпечення лексикографічної об'єктивності.* Для втілення цього принципу необхідно було з вихідного тексту СУМ-11 видалити цитати з творів марксистсько-ленінського спрямування, зняти «ідеологічну» лексику, яка в наш час уже втратила сенс (про що йшлося вище), очистити тлумачення від ідеологічних нащарувань

попередньої епохи. Крім того, лексикографам важливо було при укладанні словникових статей, формулюванні дефініцій максимально об'єктивно підходити до фактичного матеріалу, намагатися якомога точніше дотримуватися його контекстної основи, мінімізуючи при цьому вплив особистих лінгвістичних та когнітивних уподобань.

3. *Істотне розширення текстово-ілюстративної бази Словника:*

- максимально повне представлення контекстної основи, лексики та фразеології Святого Письма;
- істотне розширення списку цитованих джерел за рахунок залучення творів раніше заборонених письменників і письменників нової доби;
- залучення нових текстово-ілюстративних джерел – Інтернету й створеного в Українському мовно-інформаційному фонді лінгвістичного корпусу.

4. *Відтворення питомих рис української мови та українського менталітету:*

- введення оригінальних мовних конструкцій, притаманних творчості раніше заборонених письменників, автентичних лінгвістичних джерел й іншого фактичного матеріалу з української етнографії та фольклору;
- залучення лексикографічного матеріалу, який відображає національно-історичні реалії та ментальні концепти українського народу;
- забезпечення відповідності лексикографічного матеріалу граматичним, лексичним і стилістичним нормам сучасної української літературної мови, зокрема вилучення росіянізмів, невдалих кальок та ін.

5. *Відбиття лексико-семантичної динаміки, яка відображає сучасну мовну практику України:*

- становлення України як незалежної держави; зміна суспільного ладу; демократизація суспільства; зміна політичної ситуації в країні; зростання політичної та економічної активності всіх верств українського суспільства;
- науково-технічна революція; перехід до інформаційного суспільства та суспільства знань;
- глобалізація, розширення та інтенсифікація міжнародних зв'язків України.

6. *Осучаснення лексикографічного матеріалу:*

- введення до реєстру слів на позначення нових реалій, що з'явилися після публікації СУМ-11;

— вилучення лексикографічного матеріалу, який утратив актуальність і не властивий сучасній мовній практиці.

7. Заповнення лексикографічних лакун:

— введення до реєстру наявної в мові, але пропущеної в СУМ-11 загальноживаної лексики;

— урахування лексичних значень, які з різних причин не були відображені в СУМ-11;

— значне розширення фразеологічної зони та взагалі зони словосполучень.

Сформульовані принципи стали основою всієї роботи над укладанням СУМ-20; їх зміст детально розкриватиметься у подальшому викладі.

Перейдемо до розгляду того, як сформульовані принципи практично реалізовані у корпусі СУМ-20.

По-перше, після формування лексикографічної бази даних «СУМ» (цей процес детально описано в розділі 6) проведено роботу з виправлення помилок, недоглядів і похибок у структурі СУМ-11. Завдяки застосуванню створеного у Фонді інструментального комп'ютерного комплексу її виконання значною мірою вдалося автоматизувати. Зокрема, в автоматизованому режимі здійснено такі операції:

— аналіз знакової системи Словника;

— перевірка коректності шрифтового виділення структурних елементів словникових статей;

— перевірка вживання розділових знаків;

— перевірка коректності відсилай у значеннях та відтинках значень;

— перевірка коректності відсилай у фразеологічних частинах;

— перевірка замкненості лексикографічної системи СУМ-20 (наявність у реєстрі всіх слів, що містяться у словникових дефініціях);

— перевірка наявності ілюстрацій до значень;

— перевірка відповідності першоджерел затвердженому списку;

— перевірка вживання типових формул тлумачення в межах певних лексико-семантичних груп;

— перевірка застосування умовних скорочень у стандартних лексикографічних ситуаціях;

— перевірка узгодженості з базою даних електронного граматичного словника української мови, який функціонує в Українському мовно-інформаційному фонді НАН України у складі Національної словникової бази.

Наступний крок, де також застосовано комп'ютерні методи, — маркування ілюстрацій суспільно-політичної літератури радянського періоду та іншого заідеологізованого ілюстративного матеріалу. Результати цієї роботи значні. Наведемо лише кілька прикладів словникових статей у тому вигляді, як вони представлені у СУМ-11, підкресливши частини, вилучені в процесі подальшої роботи:

АБСОЛЮТИЗМ, у, ч. Форма керування державою, при якій верховна влада повністю належить одній особі — монархові; самодержавство. До повалення феодалізму, абсолютизму і чужонаціонального гніту не могло бути й мови про розвиток пролетарської боротьби за соціалізм (Ленін, 21, 1950, 262).

АГІТАЦІЯ, і, ж. 1. Дія за знач. агітувати. Повість [А. Головка] "Пасинки степу" пройнята глибокою вірою в перемогу колгоспів, вона була умілою і переконливою агітацією за них (Іст. укр. літ., II, 1956, 432).

2. Громадська й політична діяльність, робота серед трудящих, спрямована на роз'яснення політики якої-небудь партії чи організації. В нерозривному зв'язку з пропагандою стоїть агітація серед робітників.. (Ленін, 2, 1948, 293); Вразила мене тільки одна справа в тутешній виборчій агітації, се так звані "видатки на агітацію" (Лі. Укр., V, 1956, 41); В. І. Ленін підкреслював необхідність того, щоб пропаганда і агітація велись рідною мовою населення (Рад. Укр., 28.IX 1956, 1).

АКТИВНІСТЬ, ності, ж. 1. Властивість за знач. активний 1. Він [О. М. Сибіряков] відзначався ще надзвичайною спостережливістю, активністю та письменницьким хистом (Видатні вітч. географи., 1954, 102); В рядах комсомолу в суворій боротьбі формувалися благородні риси нашого юнацтва, його висока політична активність, невгамовна громадська ініціатива, почуття обов'язку перед Батьківщиною (Рад. Укр., 16.VI 1957, 1); // Енергійна діяльність; діяльна участь у чому-небудь; протилежне пасивність. В нинішніх умовах партія вимагає від профспілок ще більшої активності в комуністичному будівництві (Рад. Укр., 21.VI 1957, 1); Зростання активності народних мас у будівництві нового життя — закон епохи соціалізму (Програма КПРС, 1961, 13).

АРМІЯ, і, ж. 1. Сукупність усіх сухопутних, морських і повітряних збройних сил держави. У справедливих арміях доля завжди прекрасна (Гончар, I, 1954, 3); Марко розповідав про свою службу в армії, про ситуації новини, про різні випадки й зустрічі (Коз., Сальвія, 1956, 41).

Діюча армія — війська, які під час війни перебувають на фронті. Щорс, навпаки, сам просить командування послати його в діючу армію на фронт (Скл., Легенд. начдив, 1957, 19); **Радянська Армія** — назва армії Радянського

Союзу з 1946 р. Партійні організації Радянської Армії у своїй діяльності керуються Програмою і Статутом КПРС і працюють на основі інструкцій, затверджуваних Центральним Комітетом (Статут КПРС, 1961, 27); Червона Армія – назва армії Радянського Союзу до 1946 р. Українське населення Галичини з невимовною радістю зустрічало Червону Армію – армію-визволителя (Іст. УРСР, II, 1957, 191).

2. Сухопутні збройні сили...

5. кого, чого, перен. Маса людей, об'єднаних спільною ознакою, справою і т. ін. Лиш ми, робітники, ми, діти Святої армії труда, Землю будем володіти. А паразитів жде біда (Інтернаціонал); Армія пролетаріату міцніє в усіх країнах (Ленін, 13, 1949, 73); Армія юних мічурінців прагне опанувати вчення І. В. Мічуріна та його методи роботи (Юним мічур., 1955, 3).

БАНК, у, ч. 1. У СРСР – державна фінансова установа, яка здійснює кредитування народного господарства і є розрахунковим центром країни; у капіталістичних країнах – фінансова установа, що концентрує капіталовкладення, які дає в позику капіталістам з метою одержання прибутків і надприбутків. капітал переходить весь час з однієї країни до іншої. Банки, ці величезні склади капіталів, що збирають.. і розподіляють його в позику капіталістам, стають з національних міжнародними.. (Ленін, 2, 1948, 90)...

БЕЗРОБІТТЯ, я, с. 1. Стан, коли не всі можуть одержати роботу, мати постійний заробіток. Застій у промисловості [Росії] створював безробіття, яке тяжко позначалось на становищі робітничого класу (Іст. УРСР, I, 1953, 639); // Наявність безробітних. В усіх іноземних державах ми бачимо промислові кризи і величезне безробіття (Ленін, 33, 1951, 91)...

Зрозуміло, що в новому Словнику подання такого лексикографічного матеріалу є абсолютно недоречним. Отже, відповідні фрагменти словникового тексту були вилучені. Деякі словникові статті доводилося переробляти докорінно, а обсяг виявлених у такий спосіб і вилучених «ідеологічних» фрагментів сягнув приблизно 20% від усього ілюстративного матеріалу СУМ-11.

Про здійснену роботу з деідеологізації лексикографічного матеріалу свідчать наведені нижче приклади переосмислення семантики певних словникових статей (здебільшого суспільно-політичної тематики). Ці приклади подано у вигляді таблиці, ліва колонка якої представляє текст із СУМ-11, а права, відповідно, перероблений текст тієї ж словникової статті з СУМ-20:

КОМУНІЗМ, у, ч. 1. Вища суспільно-економічна формація, що приходить на зміну капіталізові й ґрунтується на усупільненні засобів виробництва; має два ступені розвитку: соціалізм і власне комунізм – безкласове суспільство, у якому здійснюється принцип: “від кожного – за його здібностями, кожному – за його потребами”. *Комунізм є найвищий ступінь розвитку соціалізму, коли люди працюють з усвідомлення необхідності працювати на загальну користь* (Ленін, 30, 1951, 178); *Революція іде, І скрізь ми Бачимо її всевітню путь, – Всі шляхи ведуть до комунізму, Всі дороги в комунізм ведуть!* (Криж., Під зорями..., 1950, 139).

2. Створена Марксом, Енгельсом і Леніним унаслідок глибокого вивчення об'єктивних законів розвитку суспільства наукова теорія й тактика пролетарської революції та побудови комуністичного суспільства. *Ідеї комунізму захоплюють серця трудящих і перетворюють світ* (Ком. Укр., 10, 1959, 9).

КОМУНІЗМ, у, ч. 1. У марксизмі – вчення про теорію і практику пролетарської революції та побудову безкласового суспільства; політика й ідеологія комуністичної партії. *Після провалу ідеї комунізму прийшла загальне розчарування в універсальних ідеях, однією з яких є, на жаль, і концепція відкритого суспільства (з наук. літ.); Велика кількість селянства, що лишалася запереким ворогом комунізму, продовжувала вперту, хоч і неузгоджену партизанську війну з більшовіцками (з навч. літ.); Таких шкід і гуртків у місті ставало все більше. І все більше виходило з них не стільки вмілих майстрів, як відмінних пропагандистів нового, початого Марксом і Енгельсом вчення про комунізм* (М. Олійник).

2. Суспільно-політична формація, яка, за марксистською теорією, ґрунтується на суспільній власності на засоби виробництва та співробітництві й взаємодопомозі вільних від класового й національного гноблення трудящих; настає слідом за капіталізмом і проходить у своєму розвитку дві стадії: нижчу (соціалізм) і вищу (повний комунізм). *Згідно з уявленнями основоположників марксизму, комунізм, як суспільна формація, передбачає наявність високорозвинених продуктивних сил, відсутність поділу на соціальні класи, скасування держави, зміну функцій і поступове відмирання грошей (з наук. літ.); Сухомлин за човна ж даяти-ме .. – За це не турбуєсь, – заспокоює Віталій .. – У нас із ним уже як при комунізмі: твоє – моє, моє – твоє...* (О. Гончар); * *Образно. Вона щедрою рукою ділилася з Пшеничкою, а той з нещасними товаришами, бо скарредою він ніколи не був. Однак цей своєрідний комунізм тривав недовго* (Д. Бузько)...

СОЦІАЛІЗМ, у, ч. 1. Перша фаза комуністичної формації, суспільний лад, що приходить на зміну капіталізму, заснований на суспільній власності на знаряддя і засоби виробництва, владі трудящих, керованих робітничим класом на чолі з марксистсько-ленінською партією. *Перехід від соціалізму до комунізму є закономірний, історично неминучий процес* (Наука..., 3, 1959, 1); *Соціалізм, планове соціалістичне господарство відкривають найширший простір для всебічного прогресу науки і техніки* (Матер. XXIV з. КПРС, 1971, 45)...

СОЦІАЛІЗМ, у, ч. 1. Суспільно-політичний лад, основою виробничих відношень якого є державна власність на засоби виробництва і як мета та ідеал висувуються принципи соціальної справедливості, свободи, рівності. *Бесіди пекучі, – немов залізо, небезпечні слова, як: соціалізм, натуралізм, дарвінізм, питання жіноче, питання робітницьке бриніли, мов бджоли* (О. Кобилянська); – *Я? Хо! Дуже здоров. Але я питаю, чи вдоволені ви страйком? – Дуже вдоволений. – Наближає до соціалізму? – Так, наближає* (В. Вланніченко); *Ніколи всміхатись. Мені треба соціалізм будувати!* (О. Довженко); *Я не визнаю пролетарського правосуддя, здійсненого за каблуками й палкою! Ось це моє кредо. І так само не визнаю соціалізму, будованого тюрмою й кулею* (І. Багрянний); *Багато з них і раніш захоплювалися ідеями соціалізму* (Б. Антоненко-Давидович)...

Уже з цих небагатьох прикладів зрозуміло, що робота з подолання ідеологічних нашарувань, властивих СУМ-11, виявилася і значною за обсягом, і важливою з огляду на необхідність послідовного проведення принципу лексикографічної об'єктивності.

4.2. Лексичний склад Словника. Критерії добору лексики та формування реєстру

Лексикографічна база даних Українського мовно-інформаційного фонду НАН України на сьогодні становить найповніше зібрання лексики української мови. Функціонуючи у складі інтегрованої системи Національної словникової бази¹, вона містить – у тій чи іншій стадії розроблення – понад 550 тисяч реєстрових одиниць. При укладанні варіанта Слов-

¹ Розпорядженням Кабінету Міністрів України від 11.02.2004 № 70-р Національну словникову базу Українського мовно-інформаційного фонду НАН України внесено до Державного реєстру наукових об'єктів, які мають статус національного надбання України.

ника, призначеного для друку, найперше питання, що виникає, — це визначення того складу реєстру (інвентарю лексичних одиниць), який доречно представити в академічному Словнику української мови.

Обсяг лексики, яка залучається до того чи іншого словника, залежить передовсім від настанов і завдань конкретної лексикографічної праці. Яким має бути словник і для кого він призначений — це ті визначальні моменти, що зумовлюють реєстр слів, дібраних до філологічного словника із лексичної скарбниці мови. Отже, тип словника мотивує вибір слів, які мають увійти до реєстру лексикографічної праці, визначає принципи і межі в доборі лексичних одиниць.

Великі тлумачні словники мають своє призначення — максимально повно представити лексику і фразеологію сучасної української літературної мови, її докладне опрацювання в семантичному, граматичному і стилістичному аспектах. Нові, комп'ютерні реалії лексикографії з її практично необмеженим потенціалом до нарощування обсягів мовного матеріалу та інтеграції лексикографічних ефектів надають чималі можливості для забезпечення цього призначення. Таким чином, СУМ-20 та його подальші версії покликані якнайширше фіксувати словниковий склад української мови, подавати детальну і різноаспектну характеристику мовних фактів. Саме в СУМ-20 втілюються ідеї інтегральної лексикографії.

Як уже зазначено, формування СУМ-20 здійснювалося на основі складу слів, що увійшли в попереднє видання — СУМ-11, адже укладання нового словника було розпочато «поверх» старого, через парсинг² його тексту та створення відповідної комп'ютерної лексикографічної бази даних. З таких початкових умов постали і завдання нової лексикографічної праці, які можна сформулювати так: а) відображення об'єктивних змін у словниковому складі та семантиці, що відбулися за час, пройде між двома виданнями; б) «вирівнювання» нового словника в прийнятих для сучасного варіанта параметрах, що означає дотримання природних для української мови лексичних пропорцій (насамперед стильових), уніфікацію однорідних лексикографічних елементів, введення до реєстру всіх слів, наявних у словникових дефініціях, — тобто послідовне проведення настанов системної лексикографії; в) усунення ідеологічних підходів у фіксації словникового складу та лексикографічному розробленні мовного матеріалу, харак-

² Поняття та процес парсингу Словника докладно описується в підрозділі 6.1.

терних для попереднього видання; г) забезпечення дериваційної повноти реєстру, введення до реєстру разом із новими словами їх дериватів.

При доборі лексики до Словника діяв принцип своєрідної фільтрації лексичного матеріалу. Він полягає в тому, що серед величезної кількості слів, які становлять словниковий склад мови, необхідно відібрати її найуживанішу, а отже, найсуттєвішу частину. Тому в практиці формування реєстру СУМ-20 застосовано статистичний критерій.

Неабияку роль у доборі лексики до Словника відіграє й авторитетність джерел, з яких залучаються лексичні одиниці. Глумачний словник національної мови будуватиметься передусім на основі художніх, публіцистичних, наукових, науково-популярних творів, мови мас-медіа тощо. Новими, порівняно із СУМ-11, джерелами стали Український національний лінгвістичний корпус та Інтернет. Вагоме значення при доборі реєстрових слів має і відчуття мови авторів Словника. Мовна інтуїція, безперечно, служить внутрішнім орієнтиром у формуванні списку лексичних одиниць.

Критеріями добору лексики до СУМ-20, виходячи із загальних принципів укладання глумачних словників, стали такі:

- узгодженість із СУМ-11 та загальносистемними настановами сучасного лексикографування (критерій системної відповідності);
- введення слів, не зафіксованих у СУМ-11, заповнення лексикографічних лакун і внесення до реєстру слів, наявних у словникових дефініціях, а також їх дериватів (лексикографічний критерій);
- урахування нових реалій, що з'явилися або активізувалися у період між двома словниками (хронологічний критерій);
- урахування частотності вживання слів (статистичний критерій);
- введення слів, зафіксованих в авторитетних джерелах, зокрема у Святому Письмі (критерій авторитетності).

Таким чином, до реєстру Словника внесено:

1) активну лексику української літературної мови, відображену в усіх різновидах художньої літератури, починаючи з кінця XVIII століття до наших днів, Біблії, народній творчості, публіцистиці, науково-популярних і наукових працях, мові засобів масової інформації та ін.;

2) терміни і номенклатурні слова, які широко вживаються у писемній та усній мові, використовуються в популярних наукових виданнях, підручниках і посібниках, пресі і т. ін. (*абсорбент, бонус, веб-сайт, вегетативний, електрон, іонізувати, морфема, неврологія, октан, увертюра*);

3) назви сучасних народів, етнічних груп і т. ін. (*албанці, англійці, батаки, білоруси, караїми, німці, уйгури*); найменування зниклих народів і

племен, які відіграли значну роль в історії та культурі (*анти, вавилоняни, в'ятичі, галли, гунни, печеніги, скифи, франки*);

4) широковживані складноскорочені слова (*ВАК, ГЕС, держадміністрація, ЄС, об'єдн.*); складні лексичні одиниці (*прем'єр-міністр, секретар-референт*);

5) продуктивні у словотворенні частини (перші та другі) складних слів (*авто..., багато..., важко..., прото..., само...; ...водний, ...граф, ...метр, ...філ*);

6) засвідчені текстово-ілюстративним матеріалом зменшені, пестливі, збільшені, зневажливі іменникові (*берізонька, вітрище, малятко, паруб'яга*) і прикметникові (*гарнесенький, здоровенний, манюнький*) найменування;

7) біблійну і релігійну лексику (*акафіст, ангел, апостол, Бог, Великдень, висвячувати, кальвінізм, преподобний*);

8) лексичні діалектизми, зафіксовані у творах багатьох письменників (*банувати, бузько, вабкий, вандри, газда, лебедіти*);

9) наявні в джерелах: пасивні дієприкметники (*адаптований, випробуваний, набраний, перероблений, розчинений*); прикметники вищого ступеня, що утворюються з відхиленням від форми звичайного ступеня (*більший – від великий, кращий – від гарний, менший – від малий*), найвищого ступеня з префіксами *щонай-, якнай-* (*щонайбільший, якнайкращий*); прислівники (*занадто, ніжки, тричі*), в тому числі й ті, що походять від прикметників (*високо, добре, по-українському, сміливо, щиро*).

До реєстру основної частини Словника не внесено:

1) вузькоспеціальні терміни, які не набули поширення в мовній практиці;

2) власні імена та прізвища людей (*Іван, Ольга, Хмельницький, Шевченко*), крім тих, які стали загальними (*геркулес, донжуан, їуда*), власні імена з міфології різних народів (*Аполлон, Велес, Венера, Зевс, Ізіда, Перун*);

3) назви країн, держав (*Англія, Венесуела, Словаччина*); географічних об'єктів – міст, морів, річок, гір, островів тощо (для них відведено два останніх томи Словника); небесних тіл – планет, комет, супутників планет тощо, крім тих, які вживаються як загальні назви (*вавилон, рубікон, содом*);

4) діалектні (обласні) слова, які не набули поширення в літературній мові, зокрема варіанти загальнолітературних слів з місцевими фонетичними відмінностями;

5) слова індивідуального творення, що трапляються у творах окремих письменників, але не закріпилися у мовній практиці;

6) прикметники вишого ступеня, які утворюються за допомогою суфікса **-іш** (*добрий – добріший, темний – темніший, холодний – холодніший*); прислівники на **-е**, утворені від цих прикметників (*добріше, найдобріше*); прикметники найвишого ступеня нормального творення з префіксом **най-** (*наймиліший, найміцніший, найновіший*);

7) дієприслівники теперішнього часу на **-ачи (-ячи), -учи (-ючи)** (*бачачи, везучи, знаючи, малячи*) і минулого часу на **-ши** (*набувши, поклавши, принісши*).

4.3. Орфографія СУМ-20

Текст СУМ-20 відповідає нормам чинного правопису 1993 року³. Порівняно з текстом СУМ-11, який дотримується норм правопису 1960 року⁴, у ньому є низка відмінностей. Це стосується відновлення в українській абетці літери **г**, зміни порядку розташування літери **ь** (м' який знак) в абетці: **ь, ю, я** замість **ю, я, ь**. Крім того, внесені деякі зміни у правопис складних слів, уживання великої літери, закінчень відмінюваних слів, слів іншомовного походження і власних назв. Правописні відмінності між СУМ-11 та СУМ-20 систематизовано у таблиці, у лівій частині якої подано фрагменти текстів правопису 1960 р., у правій – чинного.

СУМ-11

СУМ-20

Написання літери **г**

Відсутня

Літера **г** передає на письмі задньо-язиковий зімкнений приголосний як в українських словах, так і в давно запозичених і зукраїнізованих: *агрус, гава, газда, гандж, танок, татунок, гвалт, гвалтувати, гегати, гедзь, гелотати, герготати, герготіти, гигнути, гирлига, глей, гніт* (у лампі), *готель-могель, грасувати, грати* (іменник), *гратчастий, гречний, гринджоли, грунт, гудзик, гуля, джигун, дзига, дзиллик* тощо

³ Український правопис. – 4-ге вид., випр. і доп. – К.: Наук. думка, 1993. – 235 с.

⁴ Український правопис. – 2-ге вид., випр. і доп. – К.: Вид. Академії наук УРСР, 1960. – 272 с.

та похідні від них, а також у прізвищах *Галаган, Гудзь* і под.

Апостроф

Апостроф пишеться перед йотованими після губних приголосних (б, п, в, м, ф), напр.: *б'ю, п'ять, п'є, в'язи, у здоров'ї, тім'я, м'ясо, рум'яний, мереф'янський*.

Апостроф пишеться перед я, ю, є, ї після губних приголосних (б, п, в, м, ф): *б'ю, п'ять, п'є, в'язи, у здоров'ї, м'ясо, рум'яний, тім'я, мереф'янський, В'ячеслав, Стеф'юк*.

Правопис складних слів

Разом пишуться складні іменники з **пів**, друга частина яких починається з приголосного, напр.: *півколо, півмісяць, північ, півгодини, півдюжини, півкарбованця, піввідра, піврічний, півмісячний, півгодинний; півдванадцяті, півтретьої*.

Примітка. Якщо другий елемент складного слова починається з простого чи йотованого голосного, то тоді **пів** пишеться через дефіс, напр.: *пів-огірка, пів-аркуша, пів-ящика, пів-яблука*, але: *півострів, піваркуш* і под.

Разом пишуться складні іменники з першою частиною **пів-, напів-, полу-**: *піваркуша, півгодини, півдюжини, півкарбованця, півколо, півмісяць, півогірка, пів'яблука; напівавтомат, напівсон; полукілок, полумисок*.

Примітка. Перед іменниками – власними назвами **пів-** пишеться через дефіс: *пів-Європи, пів-Кисва*.

Уживання великої літери

Назви релігійних свят, постів, а також місяців, тижнів пишуться з малої букви, напр.: *різдво, покрови, великий піст, курбан-байрам; серпень, брюмер* (але: *Брюмер* у значенні «переворот 18 брюмера 1799 р.»), *неділя*; але *Петра і Павла, Маковія, Теплового Олекси, Івана Купала*.

З великої літери, але без лапок пишуться назви культових книг: *Апостол, Біблія, Євангеліє, Коран, Псалтир*, а також назви таких релігійних понять, як *Бог* (але *бог Перун* і т. ін.), *Божя Матір, Син Божий, Святий Дух* і т. ін.

З великої літери пишуться й назви релігійних свят і постів: *Благовіщення, Великдень, Івана Купала, Петра й Павла, Покрова, Різдво, Теплового Олекси, Успіння; Великий піст, Масниця, Петрівка, Пилипівка, Спасівка* тощо.

Правопис закінчень відмінюваних слів.

Іменник. Родовий відмінок.

Закінчення **-а, -я** мають назви приміщень, будівель, споруд та ін.: *хліва, коридора, позреба, сарая, бруствера, бліндажа, куреня, млина, карниза, мезоніна, тамбура, паркана*.

Примітка. Тільки з **-у** вживаються: вокзалу, гробу, даху, залу, каналу, магазину, метрополітену, молю, палацу, плоту, поверху, тину, універмагу, шинку.

Закінчення **-у** (у твердій і мішаній групі), **-ю** (у м'якій групі) мають іменники чоловічого роду на приголосний, коли вони означають назви будівель, споруд, приміщень та їх частин: *вокзалу, танку, даху, заводу, залу, замку, каналу, коридору, магазину, мезоніну, метрополітену, молю, палацу, поверху, сараю, тину, універмагу, шинку*, але (переважно з наголосом на закінченні): *бліндажа, гаража, куреня, млина, хліва*; **-а (-я)** вживається також в іменниках — назвах архітектурних деталей: *еркера, карниза, портика*; обидва закінчення — **-а (-я)** та **-у (-ю)** приймають іменники: *мосту й моста, паркану й паркана, плоту й плота*.

Правопис слів іншомовного походження

Пишеться після всіх приголосних у географічних назвах і взагалі у власних назвах (а також у словах, що походять від них) перед дальшим приголосним і в кінці слова, напр.: *Алжир, Бразилія, Вашингтон, Еоінбург, Замбезі, Капрі, Крит, Мадрид, Чикаго, Анрі, Білло, Гальвані, Грім, Дідро, Дізел, Овідій, Россіні, Тіціан, Ціцерон, Міллер; алжирський, корсіканець, сіцілієць, шіллерівський*.

Примітка. З **и** пишуться такі географічні назви: з кінцевим **-ика** — *Америка, Арктика, Антарктика, Африка, Балтика*; чотири давно узвичаєні назви із сполученням **-ри** — *Великобританія, Париж, Рига, Рим*, а також *Єгипет, Єрусалим, Палестина*.

И пишеться:

1. У географічних назвах з кінцевими **-ида, -ика**: *Антарктида, Атлантида, Флорида, Адріатика, Америка, Антарктика, Арктика, Атлантика, Африка, Балтика, Корсика, Мексика*.

2. У географічних назвах після приголосних **дж, ж, ч, ш, щ** і **ц** перед приголосним: *Алжир, Вашингтон, Вірджинія, Гемпшир, Жиронда, Йоркшир, Лейпциг, Циндао, Чикаго, Чилі*, але перед голосним і в кінці слова пишеться **ї**: *Віші, Шюфок*.

3. У географічних назвах із звукосполученням **-ри-** перед приголосним (крім **й**): *Великобританія, Крит, Мавританія, Мадрид, Париж, Рига, Рим* та ін.: але *Австрія, Рю-де-Жанейро*.

4. У раді інших географічних назв після приголосних **д, т** і в деяких випад-

ках згідно з традиційною вимовою: *Аргентина, Братислава, Бразилія, Ватикан, Единбург, Єгипет, Єрусалим, Китай, Кордильєри, Пакистан, Палестина, Сардинія, Сиракузи, Сирія, Сицилія, Скандинавія, Тибет* і в похідних від них: *аргентинець, аргентинський* та ін.

У словах, запозичених з французької мови, після шиплячого *ш* пишеться відповідно до французького *и* українське *у*, а не *ю*: *брошура, парашут*.

І пишеться після голосного: *мозаїка, наїєний, прозаїк, руїна, тейн; Енеїда, Ізмаїл*.

Але в складних словах, де перша частина закінчується голосним, на початку другої частини пишеться *і*: *староіндійський, новоірландський*; так само в позиції після префікса, що закінчується на голосний, напр.: *поінформувати, доісторичний, антиісторичний*.

І, у (а також *и* німецького дифтонга *eu*) в позиції між двома голосними (в іноземній мові) в загальних назвах звичайно не передаються окремим знаком: *буєр, лояльний, пляєда, роля, саквоаяс, фаянс*. Лише в окремих словах вони передаються через *й*, напр.: *конвейєр, фейєрверк, фойє; Фейєрбах, Савоїя, Вайян-Кутор'є, Гоїя*.

У словах, запозичених із французької мови, після шиплячих *ж, ш* пишеться відповідно до французького *и* українське *у*, а не *ю*: *брошура, жури, парашут*; також у словах *парфуми, парфумерія*.

І пишеться після голосного: *мозаїка, наїєний, прозаїк, руїна, тейн; Енеїда, Ізмаїл, Каїр*.

Але в складних словах, де перша частина закінчується голосним, на початку другої частини пишеться *і*: *староіндійський, новоірландський*; так само в позиції після префікса, що закінчується на голосний або приголосний: *антиісторичний, доісторичний, поінформувати; безідейний, дезінтеграція, дезінфекція, дезінформація, розіграш*.

І, у (а також *и* німецького дифтонга *eu*) в позиції між двома голосними (в іноземній мові) в загальних назвах звичайно не передаються окремим знаком: *буєр, конвеєр, лояльний, параноя, пляєда, роля, саквоаяс, секвоя, фаянс, фєєрверк*, але: *Гоїя, Савоїя, Фейєрбах*; також *маїя* (народність), *фойє*.

У загальних (невласних) назвах іншомовного походження приголосні не подвоюються, отже, пишеться: *аккумуляція, белетристика, грип, група, ідилічний, інтелектуальний, колектив, комісія, комуна, сума, шофер та ін.*

Тільки в небагатьох загальних іменах зберігається подвоєння н: *аннали, бонна, беладонна, ванна, манна, тонна, фини (але фінка); подвоєння л у словах білля, вілла, булла і подвоєння р у назві рослини дурра та в назві рослини й ароматної смоли мірра.*

Примітка. У невідміюваних словах італійського походження подвоєні приголосні зберігаються, напр.: *барокко, інтермецо, лібретто, нетто, фортіссімо, стаккато, брутто.*

Правопис власних назв

Літера *є* пишеться у прізвищах на *-єв*, якщо голосною не передують шиплячі або *ц*, напр.: *Голубєв, Пєвєєв, М'ятлєв*; але: *Нєхорєєв, Под'єчєв, Усвятчєв*, а також *Ботєв* (болг.), *Друмєв* (болг.).

Географічні назви

Через дефіс пишуться складні прикметники, що є частиною власних географічних назв і починаються із слів *східно-, західно-, північно-, південно-*, напр.: *Східно-Китайське море, Західно-Казахстанська область, Північно-Осетинська АРСР, також Південно-Подільське.*

У загальних назвах іншомовного походження приголосні звичайно не подвоюються: *аккумуляція, барокко, беладона, белетристика, бравісимо, ват (хоч Ватт), грип, група, групето, гун (гуни), ідилічний, інтелектуальний, інтермецо, колектив, комісія, комуна, лібрето, піанісимо, пічкато, стакато, сума, фін (фіни), фортісимо, шасі, шофер та ін.*

Тільки в окремих загальних назвах зберігається подвоєння приголосних: *аннали, бонна, брутто, ванна (ванний), мадонна, манна (манний), мотто, нетто, нанна, пенні, тонна, білля, булла, вілла, мулла, дурра, мірра.*

Російський звук *є* передається літерою *є* після приголосних (за винятком шиплячих, *р* і *ц*) у суфіксах *-єв, -єєв* російських прізвищ: *Ломтєєв, Медведєєв, М'ятлєєв, Тимірязєєв, Алексєєєв, Ведєєєєв, Гордєєєв, Менделєєєв*, але: *Муромчєєв, Нєхорєєєєв, Писарєєєв, Под'єчєєєв, Усвятчєєєв; Андрєєєєв, Аракчєєєєв, Мацєєєєв, Плєщєєєєв*, а також у болгарських: *Ботєєєєєв, Друмєєєєєв* і под.

Разом пишуться географічні назви-іменники, що складаються з прикметника та іменника, з'єднаних сполучним звуком, а також прикметники, що походять від них: *Білорілля, Верхньодніпровськ, Гостролуччя, Дорогобуж, Кам'яногірка, Красноводськ, Мажарославець, Нижньокилимськ, Нововолинськ, Новосибірськ, Старокос-*

тянтинів, Чистоводне, Ясногородка, білопільський, верхньодніпровський, гостролучанський, дорогобuzький, ка-м'яногірківський, чистоводнівський та ін.

Примітка. Це правило поширюється й на складні прикметники, утворені з двох прикметникових основ, які означають назви морів, лісів, низин, областей, залізниць тощо й мають при собі номенклатурні терміни типу *море, гора, хребет, ліс, озеро, низовина, острів* і т. ін.: *Західносибірська низовина, Нижньодунайська низовина, Новосибірські острови, Південноукраїнський канал, Східноєвропейська рівнина, Східноказахстанська область, Східнокитайське море* тощо.

5.1. Структура словникової статті

Словник української мови у 20 томах є лексикографічною системою глумачного типу. Його текст являє собою послідовність окремих елементів, кожен із яких несе змістове навантаження і має однозначне формальне відображення через набір маркерів лексикографічної системи Словника. До них належать: спеціальні символи, функціонально закріплені розділові знаки, тип шрифту тощо. Окремим видом маркера є позиційне розміщення елементів у словниковій статті. Роль маркерів дуже важлива, оскільки вони здійснюють атрибуцію елементів Словника і забезпечують цілісність його лексикографічної системи.

Основна структурна одиниця Словника – словникова стаття, що починається окремим рядком з абзацу. Кожна словникова стаття містить реєстрову одиницю – заголовне слово, яке є ідентифікатором словникової статті. Текст Словника становить послідовність словникових статей, розташованих згідно з алфавітним принципом лексикографічного впорядкування реєстрових одиниць.

Різні заголовні слова ідентифікують різні словникові статті. Однакові за написанням і наголошенням, але різні за значенням заголовні слова певних словникових статей є омонімами і супроводжуються маркером омонімії. Це верхній індекс, розташований безпосередньо після заголовного слова. – арабська цифра, якою послідовно нумеруються омоніми. Так, реєстрова одиниця СВІТ² складається із заголовного слова СВІТ і маркера омонімії – цифри «2». Лексема БАБКА має показники омонімії від «1» до «10» і з цими маркерами ідентифікує десять різних словникових статей.

Словник містить статті з одним, двома або більше реєстровими словами, об'єднаними тотожним набором лексичних значень, які утворюють реєстровий ряд. Перша лексема реєстрового ряду є заголовним словом словникової статті. Основний системно-структурний принцип побудови словникових статей полягає у спільності лексичної семантики всіх членів реєстрового ряду.

Реєстровими одиницями Словника є:

- а) лексичні одиниці;
- б) твірні компоненти складних слів: **АВТО**¹..., **ЯСНО**..., ...**ПОДІБНИЙ**.

в) слова, що вживаються лише як компоненти сталих словосполучень:

БАЗЕДОВИЙ, а, е: Δ *Базедова хвороба* *див.* хвороба.

У структурі кожної словникової статті виділяється два обов'язкових елементи: реєстрова (ліва) і тлумачна (права) частини.

5.2. Будова лівої частини словникової статті. Лексикографічний опис реєстрового ряду

У СУМ-20 представлено словникові статті двох типів: повноструктурні та відсильні. Обов'язковим елементом обох типів словникових статей є реєстровий ряд.

Ліва частина повноструктурної словникової статті містить такі елементи: реєстровий ряд (із заголовним словом); показники словозміни для змінюваних слів; показники синтаксичних зв'язків і функцій; граматичні параметри, стилістичні та інші ремарки. Загальна формальна будова лівої частини повноструктурної словникової статті з однослівним реєстровим рядом має такий вигляд:

X, Пар (**X**), Грам (**X**), Рем (**X**), Синт (**X**),

де **X** – реєстрова одиниця; Пар (**X**) – показники словозмінної парадигми змінюваних слів; Грам (**X**) – граматичні параметри (рід для іменників, вид для дієслова, показник частини мови, який подається при займенниках, числівниках, прислівниках, службових частинах мови та вигуках); Рем (**X**) – ремарки (їх докладний опис подано у підрозділі 5.8); Синт (**X**) – показники синтаксичних зв'язків і функцій.

Змінювані слова подаються в їхній початковій граматичній формі, потім наводяться показники словозміни Пар (**X**). Усі слова – компоненти реєстрових рядів, крім односкладових, подаються з наголосами – одним (але, державний) або двома (завжди, помилка). Наголоси позначаються

й у параметрах словозміни: **ВІН**, його...; **ДАВНИНА**, її...; **ПРА́ТИ**, перу́, переш...

Докладний опис способів та особливостей подання елементів Пар (X), Грам (X), Синт (X) у межах частин мови подано у підрозділі 5.4.

Кількість і склад лексичних компонентів реєстрового ряду визначається варіативністю (фонетичною, морфологічною, словотвірною, акцентуаційною, граматичною і т. ін.) заголовного слова, що не веде до зміни лексичних значень членів реєстрового ряду. Наприклад:

БАТЬКІВЩИНА¹ / **БАТЬКІВЩИ́НА**, її...; **ВІДНОВЛЮВАТИ**, юю, юєш, **ВІДНОВЛЯТИ**, яю, яєш...; **ГУРКОТА́ТИ**, очу́, бчеш, **ГУРКОТІ́ТИ**, очу́, отиш...; **ЗАВОРКОТА́ТИ**, **ЗАВУРКОТА́ТИ**, кочу́, кчеш, **ЗАВОРКОТІ́ТИ**, **ЗАВУРКОТІ́ТИ**, очу́, отиш, *док. ...*; **ЗОКРЕ́МА**...; **НАДБИРА́ТИ**, аю, аєш, *недок.*, **НАДІБРА́ТИ**, надберу́, надбереш, *док., що...*; **РОЗБОВТУВАТИ**, ую, уєш, *недок.*, **РОЗБОВТА́ТИ**, аю, аєш, *док., що...*; **СИРОТА́**, її, ч. і ж...; **УБРА́НИЙ (ВБРА́НИЙ)**, *діал.*, **УБРА́ТИЙ (ВБРА́ТИЙ)**, а, е...

Варіанти заголовного слова подаються за алфавітом, якщо вони мають однаковий стилістичний статус (**ЖИТТЄ́ВИЙ**, а, е, **ЖИТТЬО́ВИЙ**, а, е...), або за вживаністю (**МЕНШЕ**, *рідко* **МЕНШ...**).

Розроблення варіантів дієслів реєстрового ряду та похідних від них, що починаються на в та у, зумовлюється їх семантикою. Дієслова, які виражають спрямування дії в середину, в межі чого-небудь, та відповідні похідні подаються на в (із зазначенням одночасно в дужках їх варіантів на у): **ВГВІ́НЧУВАТИ (УГВІ́НЧУВАТИ)**, ую, уєш, *недок.*, **ВГВІ́НТИ́ТИ (УГВІ́НТИ́ТИ)**, нчу́, нтиш, *док., що...*; **ВГВІ́НЧУВА́ННЯ (УГВІ́НЧУВА́ННЯ)**, я, с...; **ВГВІ́НЧУВА́НИЙ (УГВІ́НЧУВА́НИЙ)**, а, е... Дієслова, які виражають поширення дії на поверхню чого-небудь, наслідок дії, та похідні від них відсилаються на у, де подаються і відповідно розробляються з варіантом в: **ВКРИВА́ТИ** *див.* **укрива́ти**; **ВКРИ́ТИЙ** *див.* **укри́тий**; **ВКРИВА́ННЯ** *див.* **укрива́ння**; **ВКРИ́ТТЯ** *див.* **укри́ття**.

Ці правила поширюються і на випадки омонімії: **ВВ'Я́ЗУВАТИ**¹ (**УВ'Я́ЗУВАТИ**), ую, уєш... 1. Вилітати в'язанням... 2. Те саме, що вилітувати 2...; але **ВВ'Я́ЗУВАТИ**² *див.* **ув'язувати**².

Применник в і прислівники з префіксальним в теж подаються на в з варіантом у: **В² (У)**, **ВГО́РІ (УГО́РІ)**; **ВТРО́Є (УТРО́Є)**.

Із слів, утворених за допомогою префікса від- та його варіанта од-, рідко вживаного в сучасній українській літературній мові, в реєстрі Словника подані тільки ті, що починаються з від-. Варіанти з од- наводяться в ілюстративному матеріалі.

Структурні варіанти подаються в окремих словникових статтях (АВОКАДНИЙ, а, е. Те саме, що авокадовий...; АВОКАДОВИЙ, а, е. Прикм. до авокадо...; БАЙДУЖІСТЬ, жості, жс. ... Стан і властивість за знач. байдужий 1...; БАЙДУЖНІСТЬ, ності, жс. Стан і властивість за знач. байдужний...).

Відсильні статті у Словнику мають таку формальну будову: *X див. у, де* *X* – заголовне слово, яке є єдиним компонентом лівої частини, а *у* – заголовне слово словникової статті, до якої йде відсилання, наприклад: *ІЗСИХАТИ див. зсихати; УНАСЛІДОК див. внаслідок.*

5.3. Будова правої частини словникової статті. Тлумачення реєстрових одиниць

Основним структурним елементом правої частини словникової статті є рубрика. У Словнику розрізняються рубрики двох типів.

Рубрика першого типу подає лексикографічне опрацювання лексичного значення. Кожна така рубрика містить тлумачення (словникову дефініцію) лексичного значення та його відтінків, а також ілюстрації до них. У полісемічній словниковій статті рубрики послідовно нумеруються арабськими цифрами. Так, словникові статті з реєстровими одиницями **З²** та **ПО** містять по 36 рубрик, **В²** – 34 рубрики.

У рубриці другого типу представлено лексикографічне опрацювання словосполучень. До неї входить текст словосполучення (реєстрова частина рубрики), тлумачення їх значень та ілюстрації до них. Докладно лексикографічне розроблення словосполучень подано у підрозділі 5.5.

Тлумачення (дефініція) як репрезентант лексичного значення у тлумачному словнику є найважливішим компонентом словникової статті. Дати слову тлумачення – це означає побудувати його семантично еквівалентний перифраз, який виступає у формі тексту, складеного відповідно до лексичних і синтаксичних правил мови, із вказівкою на елементи, що дають можливість відрізнити (упізнати, зрозуміти) означуване слово в конкретній мовній ситуації.

Словникове тлумачення і словникова стаття загалом у зручній та економній формі дають строгу і якомога вичерпнішу інформацію про значення окремого слова і показують його місце в лексичній системі, зв'язки з близькими за значенням лексичними одиницями, служать основою семантичної взаємодії цієї лексики з іншими в контексті. Утворення тлумачень базується на специфіці внутрішньої структури лексичного значення і зумовлює існування певних норм та способів їхньої побудови, що регулюються правилами, виробленими лексикографічною традицією.

Завдання дефініції у глумачному словнику формулюється так: «Глумачення мають коротко й чітко розкривати значення слова та його основні відтінки і бути по змозі однотипними для слів кожної граматичної категорії» [210:14]. Дефініція лексичного значення слова зумовлена звичайно вимогами оптимальності: вона містить такі семантичні компоненти, набір яких є необхідним і достатнім для того, щоб ідентифікувати певне значення слова і відмежувати одне значення слова від іншого.

Якщо слово має кілька значень, то на першому місці у Словнику звичайно подається значення, найуживаніше в мовній практиці. Переносне значення глумачиться, як правило, після того прямого значення, з якого воно розвинулося:

БЛІДИЙ... 1. Без рум'янцю; позбавлений природного кольору (про обличчя)... 2. Неяскравий, слабо забарвлений... 3. *перен.* Невиразний, недосконалий... **БЛИСКАТИ...** 1. *тільки 3 ос.* Раз у раз яскраво блищати, світитися переливчастим світлом; поблискувати... 2. *чим.* Уривчасто, з перервами світити, блимати (у 2 знач.)... 3. *тільки 3 ос.; перен.* Раптово з'являтися...

Якщо слово поряд із його загальноживаним значенням використовується також як термін, то загальноживане значення подається першим:

АБОРИГЕН... 1. Корінний житель країни або місцевості; тубілець... 2. *біол.* Те саме, що *автохтон* 2.; **ІНЦЮВАТИ...** 1. Сприяти виникненню чи розвиткові чого-небудь... 2. *спец.* Започаткувати, викликати хімічну або ядерну реакцію шляхом зовнішнього впливу...

Термінологічне значення подається першим лише тоді, коли загальномовне значення є перенесенням або розширенням термінологічного значення:

АДРЕНАЛІН... 1. *фізіол.* Гормон мозкового шару надниркових, що виконує важливі функції в життєдіяльності організму людини і тварин... 2. *фарм.* Лікувальний препарат, який діє *перев.* на периферійні нейромедіаторні процеси організму... 3. *перен.* Те, що викликає сильне збудження, хвилювання, страх і т. ін.; **АПОГЕЙ...** 1. *астр.* Найвіддаленіша від центру Землі точка орбіти Місяця, штучного супутника Землі або ракети... 2. *перен.* Найвищий ступінь, розквіт, вершина чого-небудь...

Основний спосіб побудови глумачення реалізується через використання описової конструкції, за допомогою якої розкривається співвідношення між словом і явищем позамовної дійсності, яке воно позначає:

АБАЖУР... Частина світильника, звичайно у вигляді ковпака, призначена для зосередження і відбиття світла та захисту очей від його впливу; **ВИПРЯМЛЯТИ...** Робити прямим, розправляти або ставити прямо, рівно що-небудь зігнуте, вгнуте, вигнуте і т. ін.; **ЦІЛЕСПРЯМОВАНИЙ...** 1. Який прагне до певної мети... 2. Спрямований на досягнення певної мети...

Службові частини мови та вигукі, зважаючи на особливості їхніх лексичних значень, тлумачаться за допомогою специфічних формул. Зокрема, тлумачення приєменників починаються словами «Уживається на позначення...», «Уживається при вказуванні на...»:

ВНАСЛІДОК (УНАСЛІДОК)... Уживається при вказуванні на причину, зумовленість чого-небудь. **ПОБІЛЯ**. 1. ... Уживається на позначення місця, предмета, особи, недалеко від якого (якої) відбувається дія або перебуває хто-що-небудь; біля, коло...

У формулах тлумачення сполучників використовуються переважно дієслова третьої особи теперішнього часу: «Поєднає...», «З'єднає...», «Приєднає...», «Починає...», «Виступає на початку...» та ін.:

АЖ²... 1. ... Приєднає підрядне наслідкове речення до головного; так що...

Тлумачення часток починаються словами «Уживається для творення...», «Уживається для вираження...», «Уживається для виділення...», «Уживається для підкреслення...», «Уживається для підсилення...», «Уживається при вказуванні...», «Уживається на початку / в кінці речення...»:

БИ... 1. Уживається для творення дієслівних форм умовного способу...

При тлумаченні вигуків переважно використовуються конструкції: «Уживається для вираження...», «Уживається для підкликання...», «Уживається при звертанні...», «Уживається як...» та ін.:

БА¹... Уживається для вираження здивування, здогаду, дивись...

Для опису звуконаслідувань використовуються формули тлумачення, які починаються словами «Звуконаслідування, що відтворює...», «Звуконаслідування, що означає...»:

БЕ¹... Звуконаслідування, що відтворює крик кози або вівці...

До описових тлумачень реєстрових слів можуть подаватися (після крапки з комою) їх синоніми, якщо вони семантично рівнозначні й загальноживані:

АГЕНТ... 3. Таємний співробітник органів розвідки якої-небудь держави, організації, шпигун...; **БЛАГОРОДНИЙ**... 1. Який відзначається високими моральними якостями; чесний, порядний...

Крім синонімів, у тлумаченнях можуть подаватися (так само після крапки з комою) антоніми:

БАГАТО... 2. *присл.* У великій кількості; протилежне мало...; **ГУСТІЙ**... 1. Який складається з великої кількості близько розміщених один біля одного однорідних предметів, частин і т. ін.; протилежне рідкий...

Описовий тип тлумачення, як основний у семантичній характеристиці слів тлумачного словника, зрештою, визначає пояснення всіх слів, що входять до реєстру, оскільки останні завжди тлумачаться через прямі чи опосередковані відсилання на слова, які мають описове тлумачення.

Іншим типом тлумачення є відсилання до синонімічного слова, тотожного певному слову за предметно-поняттєвою належністю. Літературний синонім виступає у правій частині до: а) застарілих реєстрових слів, зокрема стилістично забарвлених:

АРМАТА... *заст.* 1. Гармата...; **ВАГАДЛО...** *заст.* Маятник...; **ЛАНТИ...** *книжн., заст., поет.* Щоки...; б) широковживаних діалектизмів (якщо такий синонім існує): **КУЛЯ²...** *діал.* Милиця...; **КЕРНИЦЯ, КИРНИЦЯ...** Криниця...; але **ДИМАВКА...** *діал.* Ручка, держак біля ковальського міха або коло насоса, яким качають воду з криниці.

Синонім виступає у складі формули «Те саме, що...»: **БАГАТІЙ...** Те саме, що багач... Ця формула цілком правомірно застосовується до синонімічних пар, які диференціюються за частотністю або нормативністю, наприклад: **БУСОЛ, БУСЕЛ...** Те саме, що лелєка...; **ВІСИПКА...** 1. *розм.* Те саме, що насип...

Тлумачення ширшого за предметно-поняттєвою належністю слова здійснюється шляхом перерахування найменувань кількох вужчих предметів, дій, ознак і т. ін., які входять до об'єкта його позначення:

АГРАРІЙ... 3. Хлібороб, селянин...; **ВІДОМИЙ...** 3. Загальновизнаний, установлений, визначений...; **СПРОКВОЛА...** Не поспішаючи, не кваплячись, повільно...

Перерахування складників певного поняття може подаватися як доповнення до описового тлумачення:

БІДАЦТВО... 1. Стан убогого, убогість, бідність...; **ХУДОЖНИК...** 1. Творчий працівник у галузі образотворчого мистецтва; живописець, маляр, графік, скульптор...

Значення ряду похідних слів розкривається через відсилання (за певними формулами) на основні слова, з якими вони пов'язані.

1. Віддієслівні іменники на **-ння, -ття** зі значенням дії або стану чи одночасно дії та стану тлумачаться через відповідні дієслова: **БІДУВАННЯ...** Стан за знач. бідувати...; **ВАРІННЯ...** Дія за знач. варіти 1, 2...; **КВІТУВАННЯ...** Дія і стан за знач. квітнути...

Інші значення таких іменників нумеруються окремими цифрами: **СТАРІННЯ...** 1. Дія і стан за знач. старіти... 2. *спец.* Зміна форм і властивостей речовини під впливом різних фізико-хімічних умов або з бігом часу...

Віддієслівні іменники, у значенні яких поєднуються дія та звуки, тлумачаться за такою формулою: **БРЕНЬКАННЯ...** Дія за знач. бренькати і звуки, утворені цією дією...; **КВОКТАННЯ...** Дія за знач. квоктати і звуки, утворені цією дією...

Похідні іменники на **-ість** та іменники на **-зна**, **-ота**, що утворилися від прикметників і мають значення якості, властивості, стану, тлумачаться через відповідні прикметники за формулою: **АБСОЛЮТНІСТЬ...** Абстр. ім. до абсолютний 1...; **ГУСТОТА...** Властивість і стан за знач. густий 1-3...; **СТАРАННІСТЬ...** Властивість за знач. старанний...

Похідні іменники жіночого роду, утворені від іменників чоловічого роду, тлумачаться через відповідні слова чоловічого роду за формулою: **АКРОБАТКА...** Жін. до акробат...; **ЛЕЙБОРИСТКА...** Жін. до лейборист...; **РОБІТНИЦЯ...** Жін. до робітник 1, 2, 4...

2. Префіксовані дієслова доконаного виду, що становлять видову пару з відповідними безпрефіксними дієсловами, тлумачаться за формулою: **ЗІМІТУВАТИ...** Док. до імітувати 1, 2...; **ПОБІЛІШАТИ...** Док. до білшати...

Однократні дієслова з суфіксом **-ну-** і підсилювальні дієслова з суфіксами **-ота-**, **-оті-** подаються за зразком: **БЛИСНУТИ...** 1... Однокр. до блискати і блищати...; **СТУКОТАТИ...**, **СТУКОТІТИ...** Підсил. до стукати...

3. Значення відносних прикметників розкривається за формулами: «Прикм. до...»: **АМБРАЗУРНИЙ...** Прикм. до амбразура...; **ВЕГЕТАРІАНСЬКИЙ...** Прикм. до вегетаріанство і вегетаріанець... і «Стос. до...»: **АКЦЕНТОЛОГІЧНИЙ...** Стос. до акцентології...; **БАКАЛАВРСЬКИЙ...** Стос. до бакалавра і бакалаврату...

Загальне тлумачення «Прикм. до...» може доповнюватися семантичними відтінками: «Зробл. з...», «Який добув. з...», «Який має у своєму складі...» і т. ін. Тлумачення «Стос. до...» може розширюватися за рахунок відтінків: «Признач. для...», «Вигот. з...» і т. ін.

Якісні прикметники, які розвинулися з відносних, подаються під окремими цифрами:

БАРВІНКОВИЙ... 1. Прикм. до барвінок... 2. Який має колір квітів барвінку; голубуватий...; **КОНТАКТНИЙ...** 1. *спец.* Стос. до контакту (у 1-3 знач.), пов'язаний з ним... 2. *спец.* Який працює, впливає, діє внаслідок дотику, під час дотику... 3. *розм.* Який легко вступає в контакт, без зусиль зав'язує дружні або ділові стосунки...

Уживання прикметників, як і деприкметників, у значенні іменників відображається в тлумаченні після розробки відповідного значення прикметника як відтінку після розробки відповідного значення прислівника або як окреме значення з позначкою у знач. ім.:

ВАРЕНИЙ... Який зварився... // у знач. ім. **варена**, ної, ж. Те саме, що **варену́ха**... // у знач. ім. **варене**, ного, с. Страва або різні страви, які готуються за допомогою кип'ятіння...; **РЕЄСТРІВНИЙ**... 1. Прикм. до **реєстр**... 2. у знач. ім. **реєстрівний**, вого, ч. іст. ... Те саме, що **реєстровець**...

4. Дієприкметники тлумачаться за зразками: **ЗАГРУЗЛИЙ**... Дієпр. акт. до **загрузити** і **загрузити**...; **НАДИХАЮЧИЙ**... Дієпр. акт. до **надихати**...; **ПРИГЛУШЕНИЙ**... 1. Дієпр. пас. до **приглушити**...; **ПРИГЛУШУВАНИЙ**... Дієпр. пас. до **приглушувати**...

5. Прислівники на **е**, **о**, що походять від прикметників, тлумачаться через відсилання на відповідні прикметники: **ДОБРЕ**. 1. Присл. до **добрый**...; **ШВИДКО**. Присл. до **швидкий**...

Уживання прислівників прикметникового походження у значенні предиката подається як відтінок після розробки відповідного значення прислівника або як окреме значення з позначкою у знач. пред.: **АБСУРДНО**. Присл. до **абсурдний**... // у знач. пред. ...; **БАЙДУЖЕ**. 1. Присл. до **байлужий** 1...; 2. у знач. пред. Все одно, однаково, не має значення...

6. Для так званих дієслівних вигуків, що уподібнюються дієсловам одноактного способу дії на зразок **блиснути**, **блиснути**, **гулькнути**, **моргнути**, функція присудка є переважною синтаксичною функцією, тому тлумачна частина таких вигуків подається таким чином: «Уживається як присудок за знач. ...»: **ДМУХ**... Уживається як присудок за знач. **дмухнути**...; **ЛАП**... Уживається як присудок за знач. **лапати**...

7. У Словнику відзначаються також типові випадки образного вживання слова й усталені випадки вживання його у порівняннях, супроводжувани позначками *Образно та *У порівн.

5.4. Лексикографічна характеристика частин мови

А. Іменники

Іменники подаються у Словнику в називному відмінку однини.

При відмінюваних іменниках наводиться форма родового відмінка однини і рід: ч. (чоловічий), ж. (жіночий), с. (середній): **ВІДЕОЗАПИС**, у, ч. ...; **КАРТА**, и, ж. ...; **СПАСІННЯ**, я, с. ...

При іменниках спільного роду (чоловічого й жіночого, чоловічого й середнього, жіночого й середнього, чоловічого і середнього) ставляться відповідні ремарки: **ЛАДО**, а, ч., ж. і с. ...; **СВЕКРУШИЦЕ**, а, с. і ж. ...; **СУСІДА**, и, ч. і ж. ...; **ХЛОПЧІСЬКО**, а, ч. і с. ...

Коли родовий відмінок іменника чоловічого роду має два закінчення (-а і -у), то подаються вони обидва: **МІСТ**, моста і мосту...; **ПЛІТ**, плоту і плота, ч. ...

Якщо іменник має в родовому відмінку закінчення -а або -у залежно від значення, тоді реестрове слово оформляється за таким зразком:

АЛМАЗ, ч. 1. род. у. Мінерал кристалічної будови, що блиском і твердістю перевищує всі інші мінерали...

2. род. а. Кристал цього мінералу – коштовний камінь, огранований і відшліфований певним чином; діамант...

БУФЕТ, ч. 1. род. а. Шафа для зберігання посуду, столових предметів, закусок, напоїв...

2. род. у. Заклад громадського харчування...

Невідмінювані іменники супроводжуються позначкою *невідм.*: **АТАШЕ**, *невідм.*, ч. ...; **ПАНІ**, *невідм.*, ж. ...; **ШОСЕ**, *невідм.*, с. ...

Іменники, які вживаються тільки в множині, наводяться у формі називного відмінка із вказівкою закінчення родового відмінка і супроводжуються позначкою *мн.*: **ДРІЖДЖІ**, ів, *мн.* ...; **КУРИ**, ей, *мн.* ...; **СІНИ**, ней, *мн.* ...

Іменники, частіше вживані у множині, подаються в цій формі, а в дужках наводяться форми однини з відповідним граматичним оформленням: **ВУСА**, **ВУСИ**, ів, *мн.* (*одн.* вус, *рідко* ус, а, ч.)...; **ЛИЖІ**, лиж, *мн.* (*одн.* лижа, і, жс.)...

За таким зразком оформляються й назви народів: **ФРАНЦУЗИ**, ів, *мн.* (*одн.* француз, а, ч.; французька, и, жс.)...

Іменники, які вживаються в мові тільки у складі фразеологічних зворотів, подаються в реестрі у називному відмінку, а біля них наводиться зворот: **БАЙДИКИ**, ів, *мн.*: \diamond Байдики (байди, *діал.* гандри) бити...

У Словнику відтворюються фонетичні та акцентуаційні зміни, які відбуваються при відмінюванні слів: **ВІСЬ**, всі...; **МІЛЬ**, моли...; **НЕБЕСА**, небес...; **ОЗЕРЦЕ**, озєрця...; **ОРЄЛ**, орліа...; **ПРОСТОРІНЬ**, роні...; **ТИЖДЕНЬ**, жня...

Б. Дієслова

Дієслова подаються у формі інфінітива. До особових дієслів наводяться форми 1-ї і 2-ї особи однини теперішнього часу (дієслів недоконаного виду) або майбутнього часу (дієслів доконаного виду): **АБЕТКУВАТИ**, ую, усш, *недок.*, що...; **БРІЗНУТИ**, ну, неш, *док.*...; **ОПІСУВАТИ**, ую, усш, *недок.*, **ОПИСАТИ**, опишú, опишеш, *док.* ...

Форма 3-ї особи множини наводиться у тих випадках, коли в її утворенні є відмінність від форми 2-ї особи множини: **ГУБИТИСЯ**, гублюся, губишся; *ми*, губляться...

Якщо в особових дієсловах форма 1-ї особи практично не вживається, то подаються 2-га і 3-тя особи однини: **ДУДИТИ**, дийш, дить...

При дієсловах, що вживаються тільки з множинними або збірними суб'єктами, наводяться форми третьої особи однини і першої та другої осіб множини: **ПОЗБИГАТИСЯ**, ається, аємося, аєтесь, *док.* ...; **ПОЗЛІТАТИСЯ**, ається, аємося, аєтесь, *док.* ...

При дієсловах, які практично не мають форми 1-ї та 2-ї осіб однини теперішнього (майбутнього) часу, а також при дієсловах на **-ся** середньо-зворотного стану подається тільки форма 3-ї особи однини: **СВЕРБИТИ**, бить; **КРИШИТИСЯ**, кришиться...

До безособових дієслів подаються форми 3-ї особи однини теперішнього (майбутнього) часу й ставиться позначка *безос.*: **ВЕЧОРІТИ**, іє, *недок.*, *безос.* ...; **СПАТИСЯ**, спитьсья, *недок.*, *безос.* ...

Після особових форм дієслова у тому випадку, якщо є особливості у творенні й наголошуванні, наводяться форми минулого часу (із позначкою *мин. ч.*), форми наказового способу (із позначкою *наказ. сп.*): **БІГТИ**, біжу, біжиш; *мин. ч.* біг, ла, ло...; **ЛЯГАТИ**, аю, аєш... **ЛЯГТИ**, ляжу, ляжеш; *мин. ч.* ліг, лягла, лягло; *наказ. сп.* ляж... **ТЕРТИ**, тру, треш; *мин. ч.* тер, ла, ло; *наказ. сп.* три, трить...

Скорочення форм дієслова подаються з останнього незмінного звука основи теперішнього (майбутнього) часу: **ГАРАНТУВАТИ**, ую, уєш...; **ДРИМАТИ**, аю, аєш...; **НЕСТИ**, сую, сєш...; **РАДІТИ**, ію, ієш...

Після форм словозміни дієслова позначається його вид — *недок.* (недоконаний) або *док.* (доконаний): **БАЧИТИ**..., *недок.*; **ЗАЛИШАТИ**..., *недок.*, **ЗАЛИШИТИ**..., *док.* ... При дієсловах, які мають недоконаний і доконаний вид, даються відповідно позначення: *недок.* і *док.*: **АБСОРБУВАТИ**..., *недок.* і *док.* ...

У Словнику подані вказівки на синтаксичну сполучуваність дієслова. При дієсловах, які вимагають прямих і непрямих додатків, наводяться відмінкові питання: *кого, чого* (род. в.), *кому, чому* (дав. в.), *ким, чим* (ор. в.), *на кому, на чому* (місц. в.), а також *від кого — чого; до кого, до чого, з ким, з ким — чим, за кого, за чим, за що, кому про що, на кого, на що, по чому, про що, про кого — що, проти кого — чого, у кого, у що, по чому, про що, через що, що чим* і т. ін.

Необхідні доповнення дієслова підрядними реченнями відзначаються за допомогою позначки *із спол.*: *із спол.* що; *із спол.* що, як, чи і т. ін.; *із спол.* щоб і т. ін.

Даються також вказівки на вживання дієслів з певною частиною мови або словом, групою слів: з *інфін.*, з *прійм.* в (у), об (о) і без *прійм.*; з *прійм.* в (у), об, по і без *прійм.*; з *оруд.* в: без *прійм.* і з *прійм.* над, за, про і *т. ін.* тощо; у *сполуч.* з *присл.* легко, гарно і *т. ін.*; у *сполуч.* зі *сл.* берег, острів і *т. ін.*; у *сполуч.* зі *сл.* початок, кінець, край і *т. ін.*

Якщо дієслово вживається без додатків, використовуються ремарки *без дод.* (без додатка), *без прям. дод.* (без прямого додатка): **АГІТУВАТИ...** 2. *кого, що і без прям. дод.* ...; **БАЗІКАТИ...** *що, про що і без дод.* ...; **БОРОТИСЯ**, борюся, бóрешся, *недок.* 1. *з ким і без дод.* ...

Безособове вживання особового дієслова позначається ремаркою *безос.* за двома скісними рисками:

БУРЧАТИ, чу́, чу́ш, *недок.*, *розм.* ... 2. Те саме, що **буркотати** 2; // *безос.* Утворювати переливчасті звуки низького тону (при переміщенні їжі в шлунок)... Пасивний стан дієслів розкривається в самому тлумаченні:

ДАРУВАТИСЯ, ується, *недок.* *Пас.* до **дарувати** ...;

ЕЛЕКТРИЗУВАТИСЯ, ується, *недок.* ... 2. *тільки недок.* *Пас.* до **електризувати** 1.

Видові пари дієслів, утворених суфіксальним способом або видова співвіднесеність яких формується за допомогою переміщення наголосу, наводяться в одній статті. За вихідну береться форма недоконаного виду, а дієслово доконаного виду подається в реєстрі за алфавітом з відсиланням до відповідного дієслова недоконаного виду:

МИНАТИ, аю, аєш, *недок.*, **МИНУТИ**, ну, нєш, *док.*...

МИНУТИ *див.* **мина́ти**.

Дієслова з різними структурними суфіксами подаються в одній статті:

НАПИНАТИ, аю, аєш, *недок.*, **НАПНУТИ**, **НАП'ЯСТИ**, пну, пнєш, *док.*, *що...*;

ПЕРЕРОБЛЯТИ, аю, аєш, **ПЕРЕРОБЛЮВАТИ**, юю, юєш, *недок.*, **ПЕРЕРОБИТИ**, роблю, робиш; *мн.* *перероблять*; *док.*, *що...*

В. Прикметники

Прикметники подано в Словнику у формі називного відмінка однини чоловічого роду, потім наведено закінчення жіночого і середнього роду: **БІЛИЙ**, а, е...; **ДОЧЧИН**, а, е...; **ПІЗНИЙ**, я, е...

Невідміювані прикметники супроводжуються позначкою *невідм.*: **БЕЖ**, *прикм.*, *невідм.* Те саме, що **бе́жевий**...

Якщо прикметник має коротку форму, вона наводиться після повної: **ЗЕЛЕНИЙ**, *поет.* **ЗЕ́ЛЕН**...; **ЛАДНИЙ**², **ЛА́ДЕН**...; **РА́ДИЙ**, *рідше* **РАД**...

Вищий ступінь порівняння прикметників, утворений від іншої основи або зі зміною приголосних основи, наводиться як додаткова форма при відповідних прикметниках звичайного ступеня: **ГАРНИЙ**... вищ. ст. **кращий і ліпший, рідко гарніший**... Такі додаткові форми внесено в реєстр за алфавітом і з відсиланням до основної форми: **КРАЩИЙ**... Вищ. ст. до **гάρний 1-5, добрий 1-6, 8 і хорóбий; ліпший**... До таких слів наведено також форми найвищого ступеня порівняння: **НАЙКРАЩИЙ**... Найвищ. ст. до **гάρний 1-5, добрий 1-6, 8 і хорóбий; найліпший**...; **ЯКНАЙКРАЩИЙ**... Найвищ. ст. до **гάρний 1-5, добрий 1-6, 8 і хорóбий**...; **ЩОНАЙКРАЩИЙ**... Найвищ. ст. до **гάρний 1-5, добрий 1-6, 8 і хорóбий**...

Субстантивовані прикметники подаються або у вихідній прикметниковій статті окремим значенням чи відтінком із позначкою у знач. ім.:

БАГАТИЙ, а, е. 1... // у знач. ім. **багáтий, того, ч. Те саме, що багáч**...; або іменниковою статтею:

АЛЬБАТРОСОВІ, вих. мн. Родина морських птахів ряду буревісникоподібних...

Г. Займенники

При займенниках після граматичних форм ставиться позначка *займ.*,значається розряд. При займенниках іменникового і числівникового відмінювання подається форма родового відмінка однини: **ВИ**, вас, *займ. особ.*...; **ХТО-НЕБУДЬ**, кого-небудь, *займ. неознач.*...; при займенниках прикметникового відмінювання наводяться закінчення жіночого та середнього роду: **ВСЯКИЙ (УСЯКИЙ)**, а, е, *займ. означ.*; **СВІЙ**, свої, свої, *займ. присв.*...

У разі наявності у займенників варіантних форм і специфічних парадигм відмінювання подаються також форми інших відмінків, чисел і родів: **НІЩО**, нічого, нічому, нічим (*з прийм. ні від чого, ні з чим, ні по чому і т. ін.*), 1. *займ. запереч.*...; **САМ**, рідше **САМИЙ**, сама, самó і самé, мн. самі, *займ. означ.*...; **ХТО!**, когó, кому, когó, ким, на ко́му, біля ко́го, *займ.*...

Г. Числівники

Кількісні числівники – власне кількісні, збірні та неозначено-кількісні числівники – подаються з відмінковими закінченнями, родовими формами (там, де вони є) і в разі потреби – з формами числа, а також супроводжуються позначками *числ. кільк.*, *числ. неознач.-кільк.*, *числ. кільк. збірн.*: **ДЕВ'ЯТЬ**, ті і тьóх, *числ. кільк.*...; **ДЕКІЛЬКА**, кох і ко́х, *числ. неознач.-кільк.*...; **ОБИДВА**, обóх, ч. і с., **ОБИДВІ**, обóх, же, *числ. кільк. збірн.*...

Порядкові числівники подаються в Словнику у формі називного відмінка однини чоловічого роду, потім наводяться закінчення жіночого і середнього роду: **ДЕСЯТИЙ**, а, е...; **ТРЬОХТИСЯЧНИЙ**, а, е...

Д. Прислівники

Прислівники супроводжуються позначками *присл.* або у *знач. присл.*: **АБИКОЛИ**, *присл.* ...; **ВНИЗУ** (**УНИЗУ**), *присл.* ...; **ПО-ХОРОШОМУ**, *присл.* ...; 3. у *знач. присл.* **амфітеатром**...; **БЕЗМІР**... 3. у *знач. присл.* Дуже, надзвичайно багато... Якісні відприкметникові й віддісприкметникові прислівники на **-о**, **-е** подаються з відсиленням до твірного слова за формулою «Присл. до...»: **АРГУМЕНТОВАНО**. Присл. до аргументований...; **ЛІПШЕ**.

1. Присл. до **ліпший**...

Вищий ступінь порівняння прислівників, утворений від іншої основи або зі зміною приголосних основи, наводиться як додаткова форма при відповідних прислівниках звичайного ступеня: **БЛИЗЬКО**. 1. Присл. до близький 6; *вищ. ст.* **близьче**... Такі додаткові форми вносяться в реєстр за алфавітом і відсиленням до основної форми: **БЛИЗЬЧЕ**, *присл.* Вищ. ст. до близько... До таких слів наводяться також форми найвищого ступеня порівняння: **НАЙБЛИЗЬЧЕ**. Присл. до найближчий...; **ЩОНАЙБЛИЗЬЧЕ**, *відсил.* Присл. до щонайближчий...; **ЯКНАЙБЛИЗЬЧЕ**, *відсил.* Присл. до якнайближчий...

Е. Прийменники

Прийменники супроводжуються позначками *прийм.* або у *знач. прийм.* Біля них наводяться відповідні відмінки: **ВНАСЛІДОК** (**УНАСЛІДОК**), *прийм.*, з *род. в.* ...; **НАД**, *рідко* **НАДІ**, **НАДО**, *прийм.*, із *знах. і оруд. в.* ...; **ВСЕРЕДИНІ** (**УСЕРЕДИНІ**)... 2. у *знач. прийм.*, з *род. в.* ...

Для більшості первинних багатозначних прийменників подано формулу «Сполучення з + **реєстровий прийменник** + виражають:» та здійснено поділ на рубрики, що описують реєстровий прийменник за типом функціонально-семантичних відношень – просторових, часових, кількісних, об'єктних, порівняльних та ін. Кожна рубрика складається з підрубрик, виділених літерами українського алфавіту а), б) і т. д. для конкретизації, уточнення або деталізованого опису вживання. Наприклад:

В¹ (**У**), *рідко* **УВ** – із *знах.*, *місц.* і *род. в.*; *ц-с.*, *уроч.* **ВО**, *рідко* **ВВІ** (**УВІ**), **ВІ** – із *знах.* і *місц. в.*; *прийм.* Сполучення з **в** (**у**) виражають:

Просторові відношення

1. *Із знах. і місц. в.* Уживається на позначення предмета, місця, простору: а) всередину якого, куди спрямована дія (*знах. в.*)...; б) у якому, де відбувається дія чи хто-що-небудь міститься, перебуває (*місц. в.*)...

Об'єктні відношення

б. *Із знах. і місц. в.* Уживається на позначення дії: а) до якої хто-небудь добувається, втручається (*знах. в.*)...; б) в якій хто-небудь бере участь (*місц. в.*)...

Якщо прийменник має одне або декілька значень і переважно виражає функціонально-семантичні відношення, що належать до одного типу, поділ на частини із зазначенням типу семантичних відношень не подається: **З-ПОЗА**, *прийм. з род. в.* Уживається при вказуванні на предмет, із протилежного або зворотного боку якого спрямовано рух, дію...

Є. Сполучники

Сполучники супроводжуються позначками *спол.* або *у знач. спол.*, до яких звичайно додається назва їх відповідного семантичного розряду: **АБИ**, *спол.* 1. умовний... 2. мети... 3. з'ясувальний... 4. допустовий...; **БО¹**, *спол.* причинний...; **А²**, *спол.*... V. еднальний...; **ЗАЛЁДВЕ**... 5. у знач. *спол.*...

Складені й парні сполучники, до складу яких входять іменники, подаються при першому своєму складнику або при слові, яке містить основне семантичне навантаження: **ЗАВДЯКІ**... ○ (1) Завдяки тому, що, у знач. *спол.*...; **ЧАС**... ○ З того часу, як у знач. *спол.*...; **ЯК³**... ○ Як..., так [і]..., у знач. *спол.*...

Якщо значення сполучника стосуються різних розрядів і в межах одного розряду виділяється кілька значень, то вводиться додаткова нумерація римськими цифрами. При цьому в межах одного розряду проводиться нумерація арабськими цифрами:

А², *спол.* I. протиставний. 1. Посднує речення, протиставні одне одному за змістом; але, проте, навпаки, та... 2. Посднує речення або члени речення, не відповідні одне одному змістом, причому зміст другого суперечить сподіваному змістові, що впливає з першого; але, проте, однак, та... 3. Посднує протилежні за змістом речення або члени речення, що мають відтінок допустовості; проте, однак, все-таки... III. *присднувальний*. 1. Присднує нові речення або члени речення при послідовному викладі думок, описі ряду предметів чи явищ...

Ж. Частки

Частки супроводжуються позначками *част.* або *у знач. част.*, до яких звичайно додається назва відповідного семантичного розряду: **ЖЕ¹**, (після голосних) **Ж**, *част. підсил.*...; **ПРОСТО**... 3. у знач. *підсил. част.*... 4. у знач. *обмеж.-виділ. част.*...

3. Вигуки

Вигуки супроводжуються позначками *виг.* або у *знач. виг.*: АХ, виг. ...; БУХ, виг. ...; ГЕТЬГЕ, виг. ...; ДМУХ, виг. ...; КАРАУЛ... 2. у *знач. виг.* ...

5.5. Подання словосполучень

У структурі словникових статей СУМ-20 виділяється зона усталених словосполучень, що наводиться після характеристик лексичних значень слова. Ця зона складається з чотирьох типів одиниць, поданих з абзацу в такій послідовності: 1) стійкі словосполучення; 2) словосполучення – еквіваленти слова; 3) термінологічні словосполучення; 4) фразеологічні одиниці. Для трьох останніх типів словосполучень використовуються відповідні позначення: кружечок (○), трикутник (△), ромб (◇).

Спільними ознаками для всіх чотирьох типів словосполучень, виділених у СУМ-20, є дві риси: стійкість, усталеність, тобто неподільність, постійність їх лексичного складу і структури (поєднання слів) та їхня відтворюваність – вживання словосполучень у процесі мовлення як готових мовних штампів. Крім того, кожен із типів словосполучень має специфічні ознаки, притаманні лише йому.

Перша група словосполучень у СУМ-20 представлена стійкими словосполученнями, тобто усталеними у своєму лексичному поєднанні зворотами, які складаються в основному зі слів у їх власних значеннях і відтворюються у процесі спілкування як неподільні мовні одиниці, в яких можлива незначна десемантизація одного з компонентів. Наприклад:

Груба їжа – проста, важка для травлення їжа...; **За алфавітом** – за алфавітним порядком літер, прийнятим у мові...; **Кишенькові гроші** – гроші на дрібні витрати...

Друга група словосполучень – еквіваленти слова, тобто своєрідні перифрастичні словосполучення, рівнозначні слову, які є цілісними усталеними відтворюваними одиницями, що прирівнюються до лексико-граматичних класів слів (частин мови) із властивими для них лексичними значеннями без десемантизації компонентів. При еквівалентах слова обов'язково подана ремарка з найменуванням частини мови, у значенні якої вони виступають. Наприклад:

○ **Більше (більш) того, що...**, у *знач. присл.* – скоріше, швидше, ймовірніше...; ○ **До безмежжя**, у *знач. присл.* – безмежно, безкінечно...

Третю групу формують термінологічні словосполучення, які стосуються різних галузей знань і є елементами відповідних галузевих терміно-систем. Наприклад:

△ **Анафілактичний шок** – тяжкий патологічний стан організму, спричинений алергією на повторне введення чужорідних білків, сироваток, медикаментів і т. ін. ...; △ **Ератичні валуни**, *геол.* – перенесені льодовиком на велику відстань валуни, які складаються переваж. із порід, відсутніх у місцях їх відкладення...

До четвертої групи словосполучень належать фразеологічні одиниці – усталені за складом і структурою лексично неподільні та відтворювані словосполучення, що виражають цілісне специфічне фразеологічне значення, яке не є сумою значень його компонентів-складників, а виникає на основі їх переосмислення переважно завдяки метафоричним і метонімічним процесам. Наприклад:

◇ **Високо нестися** – поводитися зарозуміло, погордливо, зверхньо, зневажливо ставлячись до інших...; ◇ **Підводне каміння** – прихована небезпека, перешкода при досягненні мети...

Фразеологічні одиниці завжди співвідносяться з певними частинами мови і в реченні виступають відповідними його членами. Так, відповідно до зазначених диференційних ознак, притаманних фразеологізмам, серед них виділяють 5 розрядів:

– субстантивні (іменні):

◇ **Ложка дьогтю [в бочці меду]** – що-небудь незначне, невелике, що псує якусь справу або вигляд чогось. *Отой огріх дійсно виявився ложкою дьогтю, бо прокинулась вона наступного дня з думкою про латку з бур'янів і докір Ніки (з газ.);* ◇ **Крадене весілля, заст.** – таємне одруження без згоди батьків. [Горпина Корніївна:] *Так оце у вас крадене весілля, ховане вінчання!* (І. Нечуй-Левицький);

– ад'єктивні (прикметникові):

◇ **Нечистий на руку** – схильний до злодійства, шахрайства, хабарництва і т. ін. – *Шануйтеся, хлопці. Знаю: всі ви нечисті на руку. Спіймаю, позбудетеся останньої сорочки* (М. Стельмах); ◇ **Битком набитий** ким і без дод. – переповнений. *Була Грицева стодола битком набита. Зійшлися майже цілі Тунівці* (Л. Мартович);

– дієслівні:

◇ **Ковтати слова** – говорити, читати і т. ін. нерозбірливо, нечітко. – *Як ти читаси? Чому ти ковтаси слова?* (Ю. Збанацький); ◇ **Скакати цапа** – рішуче протидіяти чому-небудь, протестувати проти чогось, заперечувати щось.

[Роман:] *Та хоч би батько цапа скакав, а я таки оженюсь на Мотрі* (Г. Карпенко-Карий);

— адвербіальні (прислівникові):

◇ *За так* – безплатно, задарма; без винагороди. – *Артисти наїхали!* – *Почому квитки?* – *Безплатно!* *Це культбригада з Христополя. За так виступатиме!* (А. М'ястківський); ◇ *Як не перерветься:* а) дуже сильно, активно. *Жайворонки як не перервуться щебечуть* (Панас Мирний); б) дуже швидко. *Фашист тікає як не перерветься* (Ю. Яновський);

— вигуків:

◇ *Так куди [ж пак]* – уживається для вираження незгоди з чим-небудь, заперечення чогось. *Шрам не раз починав говорити, так куди! Голос той так і покриває слова його* (П. Куліш); ◇ *Де в біса (в чорта)!* – уживається як заперечення, спростування у знач. ні!, де там! – *Старий! А де ти в чорта? Он віці вже в квітнику!* (Василь Шевчук).

Усі лексикографовані словосполучення отримують у словниковій статті наскрізну нумерацію¹. Розроблення всіх типів словосполучень подається в словниковій статті після всіх значень слова і становить самостійний розділ статті.

Вибір опорного компонента – реєстрового слова, біля якого наведено повну лексикографічну характеристику словосполучень, здійснюється у СУМ-20 за формально-граматичним критерієм, тобто встановленням незалежного у граматичному плані компонента, який керує всіма іншими складниками або узгоджується з ними. У статтях, реєстровим словом яких є інші повнозначні компоненти такого словосполучення, подаються відсилання на опорне слово. Наприклад, термінологічне словосполучення **△ Варикозне розширення вен** з його вичерпним лексикографічним опрацюванням (тлумаченням, ремарками та ілюстраціями) подано в статті **РОЗШИРЕННЯ**, а в статтях **ВАРИКОЗНИЙ** і **ВЕНА** наведено лише відсилання за допомогою *див.*: **△ Варикозне розширення вен** *див. розширення*; ◇ **Круто заміс місити** фіксується при реєстровому слові **МІСИТИ**, а в статтях **КРУТО**, **ЗАМІС** наведено лише відсилання.

Усі словосполучення подані у вихідній реєстровій формі. Установлення вихідної форми фразеологічних одиниць залежить від їх співвіднесеності з відповідними частинами мови, що зумовлює визначення опорного компонента. У реєстрі фразеологізми зі змінюваними опорними компонентами фіксуються у прийнятій для словників формі. Так, субстантивні

¹ У монографії фразеологізми наводяться без нумерації, оскільки в процесі роботи над Словником вона буде змінюватися.

фразеологічні одиниці, вживані в однині, подаються у називному відмінку однини (♦ Біла ворона...), а вживані лише в множині – у називному відмінку множини (♦ Авгісі стійні...). Ад'єктивні одиниці зі змінними формами прикметникових (дієприкметникових) компонентів – у формі чоловічого роду однини (♦ Нечистий на руку...), крім тих, які вживаються лише у множині й відповідно наводяться у називному відмінку множини (♦ Одним миром мазані...). Фразеологізми з незмінними формами подаються у своїй єдиній парадигмі (ПРОБА... ♦ Високої проби...; ВОДА... ♦ Чистої [чистісінької] воді...).

Дієслівні фразеологізми, які вживаються в особових формах, наводяться переважно в інфінітиві, наприклад: ♦ Заорати (заріти) носом (рідко зубами). Дієслівні за формою фразеологізми, які вживаються лише у незмінній формі, подаються не в інфінітиві, а у формі, зафіксованій у фразеологізмі, наприклад: ♦ [І] через губу не плісне – дуже гордовитий, зарозумілий...

Дієприслівникові форми дієслівних фразеологічних одиниць у реєстрових фразеологізмах не наводяться, а можуть подаватися в ілюстраціях. Це не стосується фразеологізмів з формою дієприслівника, у яких розвинулося додаткове семантичне значення. Наприклад, поряд з дієслівним фразеологічним зворотом ♦ Не покладати рук (робити, працювати дуже важко, завзято...) виділяється як самостійна фразеологічна одиниця ♦ Не покладати рук – невтомно, сумлінно, з великим завзяттям, ентузіазмом...

Дієприкметникові форми дієслівних фразеологізмів фіксуються у СУМ-20 як самостійні окремі одиниці у відповідній для них вихідній формі, з можливими предикативними формами на -но, -то: ♦ Виведений з ладу – зіпсований, непридатний для використання, роботи; поламаний, понівечений...

СУМ-20 відбиває таку специфічну ознаку усталених словосполучень (переважно фразеологічних одиниць), як варіантність їхніх компонентів, тобто їх здатність до видозміни: заміни їх іншими формами, лексемами і т. ін. при збереженні семантики та лексико-граматичної характеристики. Такі варіанти фіксуються в круглих дужках (...) біля основного, інваріантного компонента, наприклад: ♦ З усією (великою, глибокою і т. ін.) переконливістю...; ♦ Впіймати (піймати) габу (ворону)...

СУМ-20 фіксує всі можливі типові варіанти словосполучень: фонетичні, що впливають зі зміни звукового складу; формальні, які стосуються зміни форми компонента, та морфологічне варіювання, що відбиває специфіку будови компонента і передбачає зміни у межах однієї або кількох форм, а також лексичну варіацію, пов'язану зі заміною компонентів іншими лексемами. Наприклад:

◇ Впа́дати (рідше па́дати) / впа́сти в око (в очі, у вічі)... ◇ Диві́тися (поглядáти) згорі́ (звѣрху) вниз (звисока́, звѣрхньо)... ◇ Гі́ти / пі́ти до рук (у ру́ки, рідко у ру́ку)... ◇ Перетяга́ти (перетягува́ти) / перетягти (перетягну́ти) на сторо́ну (на бік)... ◇ Про (на) всяки́й (всяк) випадок... ◇ Халі́ф (калі́ф) на час (на го́дину)...

Фонетичні варіанти префіксальних, приєднаних та сполучникових компонентів, таких як у – в, і – й, з – із – зі, від – од, у реєстрових фразеологічних одиницях не наводяться, а фіксуються лише в ілюстраціях, хоч у реєстрових словах вони зазначаються, наприклад:

ВБГА́ТИ (УБГА́ТИ) і ВВІБГА́ТИ (УВІБГА́ТИ)... ◇ Вбга́ти в го́лову що – усвідомити, сприйняти, засвоїти, зрозуміти або запам'ятати що-небудь. – *Це для вас в житті небезпечна річ, – сказала Каралаєва, не примічаючи, що перекручує імення, котрі вона не могла собі вбгати в старечу голову ніяким робом (І. Нечуй-Левицький); Хома ніяк не міг убгати в голову думку, що він мусить терпіти голод, коли в скрині лежать гроші (М. Коцюбинський); ГІ́ТИ (ІТИ)...* ◇ Гі́ти / пі́ти до рук (у ру́ки, рідко у ру́ку) кому – легко діставатися, здійснюватися і т. ін. – *На лови він їздив удень і вночі, риба йому йшла до рук, як приворожена. Так він проживав – цей сліпий дід (Ю. Яновський); [Не ч и по р е н к о:] Фортунить вам, Маріє Кирилівно, просто інший раз аж не віриться. Кожна справа ніби сама в руки вам іде (О. Левада); ...Наука в Остапа Тура Гнатові Яворському пішла в руку (С. Чорнобривець).*

У СУМ-20 відбито також факультативні варіанти компонентів – такі, які не є обов'язковими, через що фразеологізм функціонує у мовленні з ними або без них. Вони виділяються у реєстровій фразеологічній одиниці квадратними дужками [...], наприклад: [Від слова] до слова...; ◇ І [в сні] не сні́ться / не присні́ться...

В одному фразеологізмі можуть поєднуватися кілька варіантів, які позначаються квадратними або круглими дужками, наприклад:

◇ [Аж (рідко білий)] сві́т кру́титься (вѣрнеться, макі́триться, колиха́ється, іде) / закру́тився (завертѣвся, замакі́трився, заколиха́вся, пішо́в) о́бертом (перекі́дом, перекі́дьки, ходо́ром і т. ін.) [в о́чах (пѣред о́чіма)]...; ◇ [І (ні, ані)] ко́пійки [зла́маної (мі́дної, шербáтої)] (ше́ляга) не-ма́ (нема́є, не було́) за душе́ю...

Домагаючись однотипного оформлення реєстрових дієслівних фразеологічних одиниць і беручи до уваги допущену у мові певну довільність порядку компонентів у їхньому складі, автори на перше місце у реєстрі ставлять дієслівний компонент. Наприклад: ◇ Грі́зти / прогрі́зти го́лову... при можливому ◇ Го́лову грі́зти / прогрі́зти...

Адвербіальні та вигуківі фразеологізми у вихідній формі подаються у своїй єдиній парадигмі, наприклад:

◇ Куди твоє (рідше ваше) діло, ірон. — уживається для вираження захоплення або здивування ким-, чим-небудь; ◇ Під гарячу руку (заст. руч). — у момент великого збудження, роздратування, тимчасового гніву.

У реєстрових словосполученнях, особливо у фразеологічних, для позначення певного варіювання вживаються кутові дужки <...>. Вони використовуються у випадках, коли:

— лексичний варіант компонента має інше число, рід, відмінок і т. ін., що веде до зміни інших компонентів, наприклад: Санітарна машина <Санітарний автобус>...; ◇ Будь здоров (здоровий, здоровенький) <Будьте здорові (здоровенькі)>...; ◇ Поробля душя <Поробис серце>...

— простежуються певні незначні зміни варіантних компонентів словосполучень, наприклад: ◇ Нехай по тому буде <Та й будь по тому> <Так то-му й бути>...

— наявне паралельне функціонування фразеологізмів як у заперечній, так і в стверджувальній формі при збереженні тотожності їх семантики, наприклад:

◇ І зуби визбираєш <І зубів не визбираєш>, перев. зі сл. дати, грубо — уживається як погроза кому-небудь дуже боляче, відчутно побити за щось. — Цить! а то як візьму кочергу, то й зуби визбираєш, — крикнула Кайдашиха й скочила з місця (І. Нечуй-Левицький); — Ти що? — аж підвівся з стільця земський. — П'яний приїхав. Я такого дам тобі агітатора, що й зубів не визбираєш (М. Стельмах);

— існує інверсійний порядок уживання компонентів, наприклад: ◇ І так і сяк <І сяк і так>...; ◇ Ні з того ні з сього <Ні з сього ні з того>...

Важливим питанням оформлення реєстрової фразеологічної одиниці є встановлення її меж. Це стосується тих фактів, коли до складу реєстрових фразеологізмів не вносяться так звані супровідні слова, які є їх обов'язковим оточенням, зберігають своє лексичне значення (через що залишаються поза реєстровою одиницею), але які обов'язково повинні фіксуватися, бо лише з ними функціонує і реалізує своє значення ця одиниця. Наприклад:

◇ У хвості: а) (чого і без дод., зі сл. ставати, стати і т. ін.) у кінці чого-небудь, позаду всіх. Отож Віталій .. став у хвості досить довгої черги (І. Муратов)...; б) (зі сл. бути, опинитися і т. ін.) серед відсталих, серед останніх. Хочеш бути серед перших — учись, інакше опинишся у хвості, відстанеш від бурхливого, щодень цікавішого життя (із журн.)...; ◇ Як свої (своїх) п'ять пальців, зі сл. знати, вивчити і т. ін. — дуже добре, досконало. Знаю я вже їх [селян] як свої п'ять пальців (В. Винниченко)...

У випадках збігу звукового складу та будови різних типів усталених словосполучень для уникнення подання їх у різних місцях лексикографіи обрали такий спосіб їхнього лексикографічного представлення. Наприклад, усталене стійке словосполучення **Виходити (виступати) / вийти (виступити) з берегів** та однойменний фразеологізм \diamond **Виходити (виступати) / вийти (виступити) з берегів**, який походить від першого словосполучення, фіксується у СУМ-20 один раз у рубриці стійких (усталених) словосполучень як полісемічна одиниця, причому другим значенням подається фразеологізм з ремаркою *перен.*, яка свідчить про переносність фразеологічного значення, що є його невід'ємною ознакою, і повинна служити засобом розрізнення двох різнотипних словосполучень:

Виходити (виступати) / вийти (виступити) з берегів: а) розливатися, затоплювати береги під час повені або сильної зливи (про річку, озеро і т. ін.). *Колісь Полтва була бурхливою річкою і не тільки крутила потоки млинів, але й виходила з берегів, затоплюючи обійстя, зриваючи хліви й хати* (Ю. Винничук);... б) (*перен.*) перевершувати міру звичайного або усталеного. *Хима не розуміє гаразд, про що читають, але вона знає, що читають про святі речі, і те читання видається їй жалібним, таким дивно жалібним, що повне Химине серце виступає з берегів* (М. Коцюбинський).

До кожного усталеного словосполучення подається його семантична характеристика – тлумачення значення. Принципи та форми тлумачення у СУМ-20 залишаються такими самими, як і в інших фразеологічних словниках.

Щодо тлумачень словосполучень, то слід зауважити, що вони не суперечать правилам тлумачення слів і підкоряються основним його вимогам, тобто вони мають бути відповідними за розміром, зрозумілими, достатніми для визначення значення, точними, чіткими та лаконічними. У формулюванні дефініцій усталених словосполучень слід уникати важких конструкцій та канцелярських мовних зворотів.

Серед усіх чотирьох типів усталених словосполучень, виділених у СУМ-20, окремої уваги заслуговують фразеологізми як особлива словосполучникова одиниця.

У розкритті фразеологічного значення доцільно враховувати його специфіку і той факт, що фразеологічні одиниці – це знаки психологічного відбиття і сприйняття дійсності. Тому основною функціональною властивістю фразеологізмів є одночасне відображення ними реалій дійсності та якісна їх характеристика. Звідси випливає, що фразеологічне значення містить у собі, крім предметно-понятійного аспекту, ще й основний, до-

датковий – конотативний, – виражений в експресивно-емоційному забарвленні.

Ці особливості фразеологічного значення відбиваються у тлумаченні. Дефініція фразеологізму, як правило, пояснює предметно-логічну частину фразеологічного значення, конотативний же елемент цього значення лише частково відбитий у Словнику переважно системою функціонально-стильових ремарок, авторських зауважень, поданих у самій дефініції. Детально про застосування функціонально-стильових позначок див. підрозділ 5.8.

У тлумаченні необхідно враховувати і граматичний аспект. «Досить нерідко словникові дефініції стають не лише переконливішими, але навіть єдино правильними, якщо за ними відчувається граматичний каркас» [113:20].

Тлумачення фразеологічних одиниць будують з урахуванням граматичного значення та їх структурно-семантичних особливостей. Характер тлумачення та його зовнішня формула вираження визначається, зокрема, закріпленням синтаксичної функції за тим чи іншим розрядом фразеологізму. Відповідно до цього слід виводити і формулу тлумачення. Наприклад, фразеологічні одиниці, які виступають:

– у ролі присудка, визначаються, в основному, дієслівною формулою тлумачення:

♦ **Не бачити смаленого вовка** – не зазнавати справжніх труднощів, випробувань у житті. *Молоді [фабзавучники] ще, зелені, не бачили смаленого вовка. Дома сидючи, загориться – поїду на шахту. Поїде, спробує, аж там не так легко, як гадалось, та й назад (Д. Ткач); ♦ Кінці [у воду] ховати і т. ін. / заховати і т. ін.* – не залишати ніяких слідів злочину, негідного вчинку. – *Мені поперед усього треба людей розумних, ... кованих на всі чотири ноги .. Щоб уміли робити моїм іменем .. і ховати кінці в воду... (І. Франко);*

– у ролі додатка або підмета – субстантивною формулою:

♦ **Вавилонське стовпотворіння, книжн.** – повне безладдя, гармидер, нестримний галас, метушня. *Голови туманіли від невирухаючого лементу цього вавилонського стовпотворіння. З ранку до ночі горляють водоноси, вищять шарманки, іржуть коні (О. Гончар); ♦ Вальпургієва ніч, книжн.* – гучна, весела гулянка. – *Доброго ранку, Саши! Як спалось? – Дякую, друже! Ніяк не спалось. А ти хіба спав? – Куди там спав! Сьогодні ж у нас вальпургієва ніч! (О. Гончар);*

– у ролі означення – ад'єктивною формулою:

♦ **Як (мов, ніби і т. ін.) з відрі (з бочки, з цебра), зі сл. іти, лити і т. ін.** а) дуже сильно, потоками (про дощ). *Дощ ллє як з бочки, мабуть, десть Бог чопок загубив та й нічим заткнуть (прислів'я); Бурхнуло з неба мов із бочки, Що промочило до сорочки (І. Котляревський); ♦ Без роду й племіні* – абсолютно

одинокій, сам-один. *Кинув він усе та й утік до Києва, ізгосм став, неприкаяною людиною, без роду й племені* (А. Хижняк);

— фразеологізми, які виконують роль обставин — адвербіальною формулою тлумачення:

♦ **Як у віночку (у вінку)** — дуже гарно, чисто, охайно. — *Гарно, як у віночку. Двоє дзеркал аж під стелю між двома вікнами!* (Л. Нечуй-Левицький); ♦ **Ні знобить ні гріє кому, безос.** — зовсім байдуже, ніскільки не турбує. *Грицько всміхнувся... — Ну, що тобі до того [як люди живуть між собою], Чітко? Скажи мені, Бога ради... Адже тобі від того ні знобить ні гріє* (Панас Мирний);

Вигуків фразеологічні одиниці, як і годиться, тлумачаться описово із застосуванням формули «Уживається для вираження...»:

♦ **Того й дивись (гляді)** — уживається для вираження можливості небажаних наслідків, явищ, подій і т. ін. ... — *Візьмеш моїх пару коней, бо їм тут не солодко. Того й дивись, що ноги протягнуть* (С. Добровольський); ♦ **Що за холера!** — уживається для вираження здивування, незадоволення; що це значить. — *Іду собі попри призьбу на причілок.. Вийшла туди, коли глип — аж ніби видіти почало. Що, думаю собі, за холера така? Ще й треті пієні не співали, а вже розвидняється* (М. Ю. Гарновський).

Особливу увагу треба звертати на тлумачення фразеологічних одиниць незамкнених структур, зокрема, щоб у формулі тлумачення зазначених форм, які є несамостійними в синтаксичному плані одиницями (що стає причиною перебудови тлумачення) на першому місці (а не в середині або в кінці самого визначення) була формула «хто-небудь...» або «що-небудь...» з неодмінним зазначенням керування (здебільшого у кого, кому). Наприклад:

♦ **Голова не половою (не солімою, не кліччям і т. ін.) набита у кого** — хто-небудь має розум. [В а с и л ь :] *Дуже вже вона маніжиться! У мене, слава Господові, голова не половою набита!* (М. Кропивницький); ♦ **Світ зав'язується / зав'язався кому** — хто-небудь стає нещасним (перев. у зв'язку з невдалим одруженням). — *Доню, Галочко! Коли так, іди за Миколу! — Добре, панотченьку, — разом сказала Галочка, а на серце мов кусок льоду впав, бачачи, що усе ближче, усе ближче зав'язується її світ* (Г. Квітка-Основ'яненко);

При поділі фразеологізмів на значення слід враховувати, що відтінки у СУМ-20 не виділяються, вони кваліфікуються як значення.

Про особливості ілюстрування усталених словосполучень див. підрозділ 5.9.

Для полегшення пошуку місця опрацювання словосполучень у СУМ-20 подано Показчик словосполучень (див. підрозділ 5.11).

5.6. Лексика і фразеологія Святого Письма

Однією із засад укладання СУМ-20 є максимально повне представлення лексики та фразеології Святого Письма. Отже, при створенні нового Словника української мови одним із найважливіших джерел поповнення його реєстру був текст Біблії, значення та авторитет якої переоцінити неможливо. Особливе місце цієї книги визначене тим, що вона є винятковою, яка підноситься над усіма іншими літературними творіннями людей, маючи особливий статус Божого одкровення як книга Божих слів, звернених до людей: «Усе Писання Богом надхнене і корисне до навчання, до докору, до направи, до виховання в праведності» (2 Тим. 3:16) [120]. Біблійні слова та значення заплановано ввести в новий Словник як пласт лексики, мінімально представлений в 11-томнику, а також у зв'язку з тим, що тлумачення багатьох понять і реалій подавалися з огляду на тодішню панівну ідеологію – атеїзм.

Для залучення біблійної лексики у 20-томник створено автоматизований конкорданс, який базується на цитованому тексті Біблії у перекладі І. Огієнка, та комп'ютерну інструментальну лексикографічну систему «Біблійний СУМ», в якій представлено біблійну лексику і фразеологію згідно з принципами СУМ-20. Це дало змогу здійснити таку роботу: ввести великий масив лексики, відсутньої в 11-томнику; змінити велику кількість дефініцій, доповнити семантичну структуру багатьох лексем; значно розширити ілюстративну базу; увести до Словника біблійні фразеологізми, яких зовсім не було в СУМ-11.

Слід відзначити, що створення «Біблійного СУМа» було непростим завданням. Біблія – це унікальний світ, де відомі нам слова, повернуті до системи біблійних тем та ідей, мають специфічні значення. Якщо семантику більшості слів мови формує найближче текстове оточення (слово-сполучення, речення), то біблеїзми є дуже місткими і багатовимірними поняттями, які набувають обертонів смислу на тлі всієї текстової структури Святого Письма, адже через них отримують мовну реалізацію провідні біблійні мотиви й теми. І тому для правильного формулювання тлумачень слів Святого Письма, крім знання його тексту, потрібна велика кількість теологічних словників, енциклопедій, досліджень із герменевтики, екзегетики, тематичних біблійних симфоній, коментарів біблійних текстів. При укладанні біблійної частини СУМа використана автоматизована система «Єврейський та грецький лексикони Стронга», де до кожного слова Старого і Нового Завіту подано варіанти перекладів. Адже без спеціальних знань такі біблійні поняття, як, наприклад, віра, спасіння, гріх, по-

каєння, святість, праведність, закон, благодать, апостол, ангел тощо правильно не витлумачиш. Наведені слова є біблеїзмами, тобто лексемами, які своїм походженням сягають тексту Біблії. Адаптовані мовною практикою, вони набули вже інших значень, що сприймаються як загальнономвні. Наприклад, ключове поняття Біблії – віра – у СУМ-20 представлено в такій спосіб:

ВІРА¹, и, ж. І. Упевненість у чомусь, у здійсненні чого-небудь. *Йшла [Соломія] уперто і завзято, з вірою, що її широкі й високі груди зламують усі переешкоди* (М. Коцюбинський); ... // *у кого*. Впевненість у позитивних якостях кого-небудь, у правильності, розумності чиєїсь поведінки. *Я завжди був великим оптимістом і дотепер не втратив віри у людей, у перемогу всього світлого над темрявою і злом* (М. Коцюбинський); // Упевненість в існуванні кого-, чого-небудь. *Віра в безсмертя. Віра в потойбічний світ. Віра в прибульців*.

2. Те саме, що довір'я. *Без міри нема віри* (Номис)...

3. *бібл.* Дар Бога людині, завдяки якому встановлюється її зв'язок з Богом. Одному бо Духом [Святим] дається слово мудрості, а другому – слово знання тим же Духом, а іншому – віра тим же Духом (Біблія. Пер. І. Огієнка); Віра від слухання, а слухання через Слово Христове (Біблія. Пер. І. Огієнка); // Визнання існування Бога. Тут довелося зачепити питання і про Бога і про чорта, про віру в першого і про всякі мудрощі другого (Панас Мирний); // Те або інше релігійне вчення, віровизнання. – Котилися і наші козачі Дурні голови, за правду. За віру Христову (Т. Шевченко)...

Слід підкреслити, що дефініції у тлумачному словнику відображають загальноновживані, загальнономвні значення слів. Завдяки наявності біблійних значень Словник, а разом з тим і мова отримують інший, вищий комунікативний рівень, на якому Бог спілкується з людьми через Своє Слово – Біблію.

Загалом робота над «Біблійним СУМом» проводилась у кількох основних напрямках:

1. Зміна дефініції слова:

СУМ-11

АПОСТОЛ, а, ч. І. Згідно з євангельськими переказами – кожний із дванадцяти учнів Христа, яких він послав проповідувати своє вчення. *Чомусь пригадалося євангельське оповідання, як апостол Петро зрікся Христа* (Тулуб, Людолови, І, 1957, 40)...

СУМ-20

АПОСТОЛ, а, ч., *бібл.* 1. (*перев. з великої літери*). Повноважний посланець, якому Ісус Христос доручив проповідувати своє вчення. *А коли настав день, покликав Він [Ісус Христос] учнів Своїх і обрав із них Дванадцятьох, яких і апостолами Він назвав* (Біблія. Пер. І. Огієнка); *Апостоли з великою силою свідчили про воскресіння Ісуса*

БЛАГОДАТЬ, і, ж. І. *ц.-с.* У релігійному уявленні — щедроги, дари, ласка невідомих, таємничих сил. *Сон, кажуть, божа благодать, — ні, часом кара божа!* (Л. Укр., І, 1951, 459); *Не слід би закладати руки та ждати з неба благодаті* (Фр., Ш, 1950, 385)...

БОГ, рідко **БІГ**, бо́га, ч. У релігійних віруваннях — уявна надприродна істота, що нібито створила світ і керує ним та вчинками людей. *Була б сестра, і був би брат. А то... нема тепер нічого. Ні бога навіть, ні пів-бога* (Шевч., II, 1953, 364)...

БРАТ, а, ч. ... 4. Член релігійного братства; чернець. [Кембль:] *Ми всі брати і сестри по Христу, і всім повинна просто "ти" казати, хоч би й самому королеві* (Л. Укр., III, 1952, 22)...

ХРИСТОС, Христа, ч. (з великої літери). Назва міфічної особи Ісуса, культ якого лежить в основі християнської релігії. *І згадував літа лихі, Погані, давнії літа, Тойді повісили Христа, Й тепер не втік би син Марії!* (Шевч., II, 1963, 40)...

Господа, і благодать велика на всіх них була! (Біблія. Пер. І. Огієнка); // Найменування Ісуса Христа (з великої літери). *Отож, святи брати, учасники небесного покликання, уважайте на Апостола й Первосвященника нашого ісповідання, Ісуса* (Біблія. Пер. І. Огієнка)...

БЛАГОДАТЬ, і, ж. І. *бібл.* Щедро милість Божа, дарована людині незалежно від її заслуг. *Закон бо через Мойсея був даний, а благодать та правда з'явилися через Ісуса Христа* (Біблія. Пер. І. Огієнка); *Спасенні ви благодаттю через віру, а це не від вас, то дар Божий, не від діл, щоб ніхто не хвалився* (Біблія. Пер. І. Огієнка); *Рушієм самовдосконалення є прагнення віруючого повернутися до стану благодаті* (із журн.)...

БОГ, рідко **БІГ**, ч. І. *род.* Бо́га, *бібл.* (з великої літери). Творець і Володар Всесвіту, усього суцього, який керує світом. *На початку Бог створив небо та землю* (Біблія. Пер. І. Огієнка); *Боже великий єдиний, Нам Україну храни. Волі і світу промінням Ти її осні* (О. Кониський)...

БРАТ, а, ч. ... 5. *бібл.* Член духовної християнської спільноти; християнин. *А ви вчителями не звіт'єся, бо один вам Учитель [Ісус Христос], а ви всі брати* (Біблія. Пер. І. Огієнка); [Кембль:] *Ми всі брати і сестри по Христу, і всім повинна просто "ти" казати, хоч би й самому королеві* (Леся Українка)...

ХРИСТОС, Христа, ч. (з великої літери). Титул Сина Божого, віра в якого лежить в основі християнської релігії; Месія. *Сьогодні в Давидовім місті народився для вас Спаситель, Який є Христос Господь* (Біблія. Пер. І. Огієнка)...

2. Уведення нових слів:

АПОСТОЛЬСТВО, а, с., *бібл.* Покликання, служіння апостола; вид посланництва від імені Ісуса Христа. *Через Нього [Ісуса Христа] прийняли ми благодать і апостольство на послух віри через Ім'я Його між усіма народами* (Біблія. Пер. І. Огієнка).

БОГОЗНЕВАГА, и, ж., *бібл.* Паплюження Бога або догматів віри. *Відкиньте і ви [віруючі] все оте: гнів, лють, злобу, богозневагу, безсоромні слова з ваших уст* (Біблія. Пер. І. Огієнка).

БОГОЗНЕВАЖАТИ, аю, аеш, *недок.*, *бібл.* Ганьбити Бога, паплюжити святе, відрікатися від Бога. *Як юдеї ж побачили наповні, то наповнились заздрощів і стали перечити мові Павла та богозневажати* (Біблія. Пер. І. Огієнка).

БОГОЗНЕВАЖНИЙ, а, е, *бібл.* Прикм. до **богозневага**. *І побачив я [апостол Іван] жінку, що сиділа на червоній звірині, переповненій іменами богозневажними, яка мала сім голів і десять рогів* (Біблія. Пер. І. Огієнка).

БОГОЗНЕВАЖНИК, а, ч., *бібл.* Той, хто ганьбить Бога, відрікається від Нього. *Будуть-бо люди тоді [останніми днями] самолюбні, грошлюбні, зарозумілі, горді, богозневажники* (Біблія. Пер. І. Огієнка).

БОГОПІЗНАННЯ, я, с., *бібл.* Пізнання Бога. *Я милості хочу, а не жертви, і богопізнання більше від цілопалень* (Біблія. Пер. І. Огієнка); *Богопізнання доступне всім охочим, але тільки тим, які вірять у Бога, духовно чистим людям* (із журн.).

3. Розширення семантичної структури слова:

БЕЗЗАКОННИЙ, а, е. 1. Який суперечить установленим законам; проти-законний. *Діяльність організованих злочинних утворень спрямована на беззаконне збагачення, створення стереотипу красивого життя, формування особливого типу "крутої" людини, здатної піти на будь-який вчинок* (із журн.); // *Який іде всупереч вимогам моралі, справедливості. Св'яті суд беззаконний над вами Катюги-судді вирікали, і ви На смерть волочили кайдани* (з народної пісні).

2. *бібл.* Який суперечить Закону Божому. *Бо досить минулого часу, коли ви чинили волю поган, ходили в розпусті, у пожадливістях, у пияцтві, у гулянках, у пиятиках, у беззаконних ідолослужбах* (Біблія. Пер. І. Огієнка); // у знач. ім. **беззаконний**, ного, ч. Той, хто порушує Закон Божий. *Закон не покладений для праведного, але для беззаконних та для неслухняних, нечестивих і грішників, безбожних та нечистих, для зневажників батька та зневажників матері, для душегубців* (Біблія. Пер. І. Огієнка).

БЕЗЗАКОННИК, а, ч., *заст.* 1. Той, хто порушує закон. *Цар... зараз же бомагу [бумагу] їй посилає, Щоб беззаконника такого-то схотить* (Л. Глібов).

2. *бібл.* Той, хто порушує Закон Божий; грішник, нечестивець. *Бо ось вороги Твої Господи, бо ось вороги Твої згинуть, розпорошаться всі беззаконники!* (Біблія. Пер. І. Огієнка); *І тоді-то з'явиться той беззаконник, що його Господь Ісус заб'є Духом уст Своїх і знищить з'явленням приходу Свого* (Біблія. Пер. І. Огієнка).

БЕЗАКОННЯ, я, с. 1. Порухення правового порядку, законів, установлених у певному суспільстві, державі, в певну епоху. *Він [Т. Шевченко] гостро говорив на судстві про пригнічення панами народу, про беззаконня, які чинили скрізь царські чиновники (П. Колесник)...*

2. бібл. Сукупність дій, пов'язаних із порушенням заповідей та повелінь Божих; гріх, неправедність, нечестя. *Ми прогрішилися та чинили беззаконня, і були несправедливі, і бунтувалися, і відверталиися від Твоїх заповідей та від постанов Твоїх (Біблія. Пер. І. Огієнка); Кожен, хто чинить гріх, чинить і беззаконня. Бо гріх — то беззаконня (Біблія. Пер. І. Огієнка); * Образно. Я бачив таких, що оралі були беззаконня та сяяли кривою, то й жали її: вони гинуть від подиху Божого і від духу гнівного Його погибають! (Біблія. Пер. І. Огієнка).*

Слід зазначити, що під час роботи з біблійною лексикою важливим процесом стало розширення тлумачного словника новими словами і значеннями, які не пов'язані з духовними поняттями, а стосуються найменувань різних реалій життя, побуту, стосунків людей, ознак, дій тощо. Це виявилось не менш потужним джерелом розвитку словникової структури, ніж поповнення біблеїзмами. Наприклад:

БАЛЬЗАМОВИЙ, а, е. Прикм. до **бальзам**. *І станеться, як ти [цар Давид] почувши шелест ніби кроків на верховіттях бальзамових дерев, тоді вийдеш на бій, бо то вийшов Бог перед тебе, щоб побити філістимський табір (Біблія. Пер. І. Огієнка)...*

БАТЬКО, а, ч. ... 3. Благодійник, покровитель. *Бідарям я був батьком, суперечку ж, якої не знав, я досліджував (Біблія. Пер. І. Огієнка)...*

БЕРЕГТИ, ежу, ежеш; мин. ч. беріг; берегла, лб; *недок.*, кого, що. ... 4. що. Дотримувати, пильнувати, виконувати. *Хто Закон береже, розумний той син, а хто водиться із гультаями, засорамлює батька свого (Біблія. Пер. І. Огієнка); По правді, по правді кажу вам. Хто слово Моє [Ісуса Христа] берегтиме, не побачить той смерті повік! (Біблія. Пер. І. Огієнка)...*

БИТИ, б'ю, б'єш; наказ. сп. бий; *недок.* ... 2. ... // *чим і без дод., перен.* Карати кого-небудь. *Я бив вас посухою та зеленячкою [борошнистою россою], сарана жерла безліч ваших садків й виноградників ваших, та ви не вернулись до Мене, — говорить Господь (Біблія. Пер. І. Огієнка)...*

БИЧ, а, ч. ... 3. *перен.* Кара, покарання. *Нехай забере Він [Бог] від мене Свойого бича, Його ж страх хай мене не жахає (Біблія. Пер. І. Огієнка)...*

БИГУН¹, а, ч. ... 3. *заст.* Слуга, який біжить попереду екіпажа, а також гонець, посылний. *А Адонія, син Хаггітши, бундючнися та говорив: Я буду царювати! І справив він собі повоза та верхівців, та п'ятдесят чоловіка [осіб] бігунів перед собою (Біблія. Пер. І. Огієнка)...*

БЛАГОВОЛІННЯ, я, с. 1. *заст.* Дія за знач. **благоволити**¹. *Хай буде над нами благовоління Господа, Бога нашого, і діло рук наших утверди нам (Біблія. Пер. І. Огієнка)...*

4. Фіксація біблійної фразеології:

◇ **Гесня вогнінна (огнінна)** – пекло. *Коли твоє око тебе спокушає, його вибери й кинь від себе: краще тобі однокім ввійти в життя, ніж з обома очима бути викиненому до гесни огненної* (Біблія. Пер. І. Огієнка)...

◇ **Дерево життя, бібл.** – за Біблією дерево в Едемському саду, плоди якого є джерелом вічного життя. *І сказав Господь Бог: Ось став чоловік, немов один із нас, щоб знати добро й зло. А тепер коли б не простяг він своєї руки і не взяв з дерева життя, і щоб він не з'їв, і не жив повік-віку* (Біблія. Пер. І. Огієнка); *Хто має вухо, хай чує, що Дух [Святий] промовляє Церквам: переможцеві дам їсти від дерева життя, яке в раю Божім* (Біблія. Пер. І. Огієнка); ◇ **Дерево пізнання (знання) добра і (й) зла, бібл.** – за Біблією, дерево в Едемському саду, плодів якого Бог заповідав не їсти першим людям Адамові та Єві. *І зростив Господь Бог із землі кожне дерево, принадне на вигляд і на їжу смачне, і дерево життя посеред раю, і дерево пізнання добра і зла* (Біблія. Пер. І. Огієнка); *І наказав Господь Бог Адамові кажучи: Із кожного дерева в раю ти можеш їсти. Але з дерева знання добра й зла не їж від нього, бо в день їди твоєї від нього ти напевно помреш!* (Біблія. Пер. І. Огієнка)...

5. Уведення ілюстративного матеріалу з Біблії, в тому числі до загальноновживаної лексики (понад 1000 ілюстрацій). Ці ілюстрації не обов'язково містять якусь релігійну настанову чи мають повчальний характер (хоча багато й таких), вони просто відбивають певні сторони дійсності. Наприклад:

АРАМЕЙСЬКИЙ, а, е. Стос. до арамеїв, належний, власт. їм. *Ісаак поблагословив Якова і послав його до Падану арамейського, щоб узав собі звіти жінку* (Біблія. Пер. І. Огієнка)...

БАГРЯНИЙ, а, е. Густо-червоний, пурпуровий. *І вийшов назовні Ісус у терновім вінку та в багрянім плащі. А Пилат до них каже: Оце Чоловік!* (Біблія. Пер. І. Огієнка); *Розжеврене, червоне сонце низько спустилося, багряним світлом градо. По деревах, по квітках і по річці* (Леся Українка)...

БАЖАТИ, аю, аєш, недок. І. чого або з інфін. Прагнути, мати бажання здійснити, одержати, здобути що-небудь; хотіти. *І як бажаєте, щоб вам люди чинили, так само чиніть їм і ви* (Біблія. Пер. І. Огієнка); *Серце моє трудне. Чого ти бажаєш, що в тебе болить?* (Т. Шевченко)...

ВИБАЧЛИВИЙ, а, е. І. Який виявляє поблажливість; несурорий. *Ти, Господи, добрий і вибачливий, і многомилостивий для всіх, хто кличе до Тебе!* (Біблія. Пер. І. Огієнка); *З таким посміхом, вибачливим, стриманим, старший дивляться на пустощі улюбленої дитини* (О. Донченко)...

ДАР, у, ь, уроч. І. Те, що підноситься, жертвується; подарунок. *Дар Божий – вічне життя в Христі Ісусі, Господі нашім!* (Біблія. Пер. І. Огієнка)...

Мова тексту Біблії, яка апелює до серця людини, є дуже образною, в ній багато епітетів, метафор, алегорій, символів, слів та виразів із пере-

посним значенням. Вона збагатила зміст СУМа багатьма порівняннями, метафоричними та фразеологічними виразами:

ЗЕРНО, зерна, с. ... 2. Насінина плодів певної рослини. *Вона навіть не вибирала її не їла зерен і викидала з рота зерна з лущипайками* (І. Нечуй-Левицький); *Особливу симпатію викликав шуктильгаючий коник Черниша... Йому давали вилущені горіхові зерна, персики, виноград, і він усе охоче поїдав* (О. Гончар); // у знач. збірн. *Зерно бобових культур має досить різноманітне використання* (з наук. літ.); * У порівн. *По правді кажу вам: коли будете ви мати віру, хоч як зерно гірчичне, і горі оцій скажете: Перейди звідси туди, то й перейде вона, і нічого не матимете неможливого!* (Біблія. Пер. І. Огієнка).

ВИТЯГАТИ, аю, аєш, **ВИТЯГУВАТИ**, ую, уєш, *недок.*, **ВИТЯГТИ**, **ВИТЯГНУТИ**, гну, гнеш; *мин. ч.* витяг, ла, ло і витягнув, нула, ло; *док.*, *кого, що...* *О* **Витягти** (**витягнути**) / **витягати** (**витягувати**) з болота (з багна, з калюжі, з грязюки і т. ін.) *кого:* а) допомогти кому-небудь вийти з принизливого становища або позбутися злиднів. *Витяг мене Він [Бог] із згубної ями, із багна болотистого, і поставив на скелі ноги мої, і змінив мої стопи* (Біблія. Пер. І. Огієнка)...

ГЛИНА, и, ж. Багатомінеральна гірська порода, перев. осадового походження, яка з водою утворює пластичну масу; широко використовується у будівництві, гончарстві, скульптурі і т. ін. *Назносим каміння, назносимо глини, Збудуємо хатку з дверима у сіни* (М. Коцюбинський); *Є добра глина, цеглу з неї роблять...* (А. Шиян); * *Образно. Тепер же, о Господи, Ти наш Отець, ми глина, а Ти наш гончар, і ми всі чин Твоєї руки!* (Біблія. Пер. І. Огієнка); * У порівн. *Усе міняється, вліплюється, мнеться, мов глина творчая у скульптора в руках* (П. Тичина)...

Таке повне і багатостороннє включення мовного матеріалу Біблії до Словника української мови дає підстави стверджувати, що цей Словник посяде особливе місце у фундаментальній національній лексикографії, оскільки тут подано об'єктивне (незаідеологізоване), докладне, різноаспектне представлення Слова Божого через систему лексикографування. Завдяки цьому здійснено реальний поступ для належної та достойної репрезентації Святого Письма у сучасній українській літературній мові.

5.7. Нова загальноживана і термінологічна лексика

Динамічні процеси в сучасній українській літературній мові протягом останніх десятиліть спричинили появу великої кількості нових загальноживаних і термінологічних слів, що потребують упорядкування та системного представлення у тлумачній лексикографії. Неологічна лексика, зафіксована у СУМ-20, розширилася за рахунок морфологічної і семантичної деривації, а також запозичень з інших мов. До неологізмів відносимо лек-

сичні одиниці, що є новими у 20-томнику порівняно з 11-томником, тобто слова і значення, які виникли в період між створенням цих двох лексикографічних систем, а також ті, які існували, але не були зафіксовані в 11-томному Словнику.

Зіставлення СУМ-11 і СУМ-20 надає можливість проаналізувати інноваційні процеси в лексико-семантичній системі української мови, виявити тенденції та шляхи її розвитку, встановити причини виникнення і характер цих змін. Суттєвим напрямком розвитку мови є інтелектуалізація, про що свідчить поява великого масиву наукової лексики, яка значно розширила реєстр СУМ-20. Утворення значної кількості лексичних одиниць різних галузей знань є результатом прискорення науково-технічного прогресу. Крім того, в сучасній мові розвивається її експресивна, образна та емоційна сторони, що демонструє постання чималого пласта нових слів на основі метафоризації.

Майже половину всіх неологізмів, утворених морфологічним способом, становлять іменники. Високою продуктивністю при цьому відзначається суфіксація. Суфіксальні утворення використовуються на позначення абстрактної ознаки, опрідметненої дії, осіб жіночої та чоловічої статі тощо.

До реєстру СУМ-20 зі значенням абстрактної ознаки (абстрактні та збірні назви) внесено велику кількість слів, де продуктивним виступає суфікс *-ість*, який стосується характеристики особи, предмета, явища, їхніх якостей, стану, властивостей. Це відприкметникові (*альтернативність, бар'єрність, бартерність, олігархічність, шикарність*) і віддієприкметникові (*деформованість, прогнозованість*) іменники. Для найменування професійної діяльності та поведінки особи, суспільних рухів тощо вживаються іменники, утворені від прикметників за допомогою суфіксів *-ств-*, *-цтв-*, (*аудиторство, байкерство, здирництво, рейдерство, угодництво*). Значне місце займають новотвори, представлені словотвірним типом із суфіксом *-ізм, -изм* (*актуалізм, аstatизм, вірилізм, глобалізм, популізм*). Указані неологізми використовуються для найменування суспільно-політичних течій, економічних процесів, методів дослідження тощо. До групи частотних слів належить також інноваційна віддієслівна лексика, утворена за допомогою суфікса *-ни-* на позначення різних процесів та дій (*адсорбування, банкрутування, видрукування, легалізування, мацерування*).

Активними серед іменникових дериватів нового Словника є також словотвірні типи для найменування осіб жіночої статі з суфіксами *-к-* (*акціонерка, банкірка, батерфляїстка, корупціонерка*), *-ниц-* (*байдарочниця,*

барахольниця, валютниця, ізолювальниця) та чоловічої статі з суфіксами **-ист, -іст** (*бар'єрист, батумист, вікіпедист, гедоніст*), **-ник** (*антенник, арбалетник, естражник*).

У поповненні реєстру Словника значну роль відіграють дериваційні неологізми, що є складними іменниковими утвореннями: *автоцивілка, бюджетоутворення, векселедержатель, ВІЛ (вірус імунodefіциту людини), експрес-служба, ЄС (Європейський Союз)* та ін. Серед цих одиниць значну групу становить наукова термінологія з фізики, інформатики, техніки й т. ін. Велика кількість слів утворена за допомогою таких компонентів: **веб-** (*веб-адміністратор, веб-адреса, веб-узел, веб-дизайн*), **гамма-** (*гамма-активність, гамма-апарат, гамма-астрономія*), **дельта-** (*дельта-деревина, дельта-електрон, дельта-проміння, дельта-частинка*), **мас-** (*мас-спектрометр, мас-спектроскопія*).

Суттєвих змін зазнала суспільно-політична та фінансово-економічна лексика, що стала одним із джерел поповнення словникового фонду української мови. Значну за обсягом групу реєстру СУМ-20 становлять іменники з компонентами: **банк-** (*банк-акцептант, банк-гарант, банк-дилер, банк-кореспондент*), **бізнес-** (*бізнес-карта, бізнес-каталог, бізнес-план, бізнес-проект*), **брутто-** (*брутто-відсотки, брутто-оренда, брутто-прибуток*), **віце-** (*віце-прем'єрство, віце-спікер, віце-спікеріада, віце-спікерство*), **медіа-** (*медіа-група, медіа-магнат, медіа-простір, медіа-центр*), **прес-** (*прес-випуск, прес-група, прес-клуб, прес-центр*), **соціал-** (*соціал-націоналізм, соціал-націоналіст, соціал-патріот, соціал-реформізм*).

Велику частину новотворів становлять складноскорочені слова. Мовний матеріал СУМ-20 засвідчує, що наймасовішим процесом поповнення реєстру є поява часткових аббревіатур. У процесі їх творення беруть участь такі словотвірні елементи: **держ-** (*держбюджет, держзамовлення, держкордон, держкредит*), **політ-** (*політапарат, політкоректність, політреформа*) та ін. Нові складні утворення збагатили також лексику військової справи (*артвогонь, артобстріл, військтехнік*). Дещо меншу частину неологізмів становить спортивна лексика (*спортклуб, спорткомітет, спорткомплекс*).

У реєстрі нового Словника є невелика група ініціальних аббревіатур із двома підтипами: літерні (утворені з початкових букв слів) та звукові (утворені з початкових звуків слів). Кількісно найбільшими серед ініціальних аббревіатур є літерні утворення (*дорожньо-транспортна пригода – ДТП, міжбібліотечний абонемент – МБА, Служба безпеки України – СБУ*), значно меншу частину становлять звукові новотвори (*засоби масової інформації – ЗМІ, Національна академія наук України – НАНУ*); мен-

шою кількістю представлені абрєвіатури, утворені з початкових частин слів. Найхарактернішими є найменування осіб за професією: *військовий кореспондент* – *військкор*, *начальник медичного закладу* – *начмед*, *помічник режисера* – *помреж*.

Вагому роль у поповненні словникового складу української мови неологічними входженнями у СУМ-20 відіграють прикметники, які за чисельністю займають друге місце після іменників. Основний їх масив – це одиниці, утворені суфіксальним способом.

Велику групу нових слів становлять відносні прикметники. У СУМ-20 виявлено значну продуктивність лексичних одиниць із суфіксами *-ичн-*, *-ічн-*. Відносні прикметники цієї групи творяться від основ іменників жіночого роду на *-ія* (*агіографічний*, *алергійний* (*алергічний*), *пиемічний*). Характерною особливістю цього словотвірного типу є значна кількість інноваційних прикметників, похідних переважно від назв наукових дисциплін та методів дослідження (*аксіологія* – *аксіологічний*, *сфігмографія* – *сфігмографічний*, *тахографія* – *тахографічний*, *фармакологія* – *фармакологічний*), а також слів із цими суфіксами на позначення релігійних понять, зокрема церковних рухів (*єкуменічний*, *свангелістичний*).

Значна частина неологізмів – відносні прикметники, утворені за допомогою суфікса *-арн-*, (*-ярн-*), мотивовані іменниковими основами чоловічого роду з кінцевим приголосним та жіночого роду на *-а* (*-я*): *атомарний* (*фіз.*), *вакуолярний* (*біол.*), *капсулярний* (*біол.*), *лейкоцитарний*, *регіонарний* (*мед.*). У межах відносних відіменникових прикметників з'явився ряд слів, утворених за допомогою суфікса *-ивн-* від іменників жіночого роду на *-ія*: *адгезивний* (*біол.*, *мед.*), *адоптивний* (*юр.*), *аплікативний* (*мат.*), *імплікативний* (*лог.*), а також похідні, співвідносні з іменниками чоловічого роду, що мають нульову флексію, наприклад: *абсолютивний*, *ад'сктивний*. Інновації такого словотвірного типу представлені переважно термінологічною лексикою з лінгвістики (*афрікативний*, *дескриптивний*, *єративний*, *імплозивний*) та біології (*адаптивний*, *адвентивний*, *гермінативний*).

Серед відносних прикметників, утворених від іменникових основ, що виражають значення загальної відносності, постала група слів із продуктивними суфіксами *-альн-*, *-іальн-*, *-уальн-*. Це лексичні одиниці на позначення лінгвістичних (*адвербіальний*, *аспектуальний*, *імперсональний*), медичних (*анестезувальний*, *аортальний*, *вакцинальний*) та інших термінів. Досить продуктивним є суфікс *-ов-*, за допомогою якого утворилися слова на позначення одиниць фінансової (*акцентовий*, *бізнесовий*, *готівковий*) та біологічної (*білковий*, *голоценовий*, *джгутиковий*) сфер.

Ряд матеріально-речовинних прикметників із суфіксом **-н-** виник від іменників чоловічого роду з нульовим закінченням та жіночого роду на **-ія**. Значна кількість цих новотворів стосується хімічної галузі (*альдегідний, гідридний, йодидний*), порівняно менша – техніки (*абзетцерний, абляційний, анероїдний*), медицини (*абсцесний, абзусний*), фінансів (*авальний, аудиторний*).

У новому Словнику зафіксовано також відносні прикметники, що мають віддієслівне походження. Це слова із суфіксом **-альн-**, що належать переважно до загальнонаукової лексики. Наприклад: *анестезувальний, балансувальний, брикетувальний, буксирувальний*.

Поповнення словникового складу української літературної мови відбувається також за рахунок якісних прикметників, котрі за кількістю значно поступаються відносним. В основному це одиниці із суфіксами **-еньк-** (*голосенький, тяженький, щедренький*), **-есеньк-** (*свіжесенький, темнесенький, холоднесенький*), **-ісінк-** (*вірнісінкий, давнісінкий, жоднісінкий*). Менше утворень із суфіксами **-усінк-**, **-юсінк-** (*маніпусінкий, худюсінкий*), що використовуються для вираження посиленого або зменшеного вияву ознаки, якості.

Вагоме місце у процесах неологізації посідає також префіксально-суфіксальний словотвір, на основі якого за допомогою префікса **без-** та суфіксів **-н-**, **-ов-**, **-льн-** утворюються насамперед прикметникові неологізми загальнонавчальної сфери (*безавансовий, безвізовий, бездотаційний, безкабельний, безлімітний*). Серед групи лексем цього типу виявлено також інноваційні одиниці, що стосуються переважно медичної (*безвітамінний, безмікробний, безопераційний*), економічної та фінансової (*безвалютний, бездотаційний*) галузей. З-поміж найвагоміших інноваційних утворень є також прикметники із префіксом **проти-** та суфіксами **-н-**, **-ичн-** (**-ічн-**), **-ськ-** (*протианемічний, противірусний, протиграмадський, протипухлинний, протирадіаційний*). Найбільше таких слів функціонує в медицині. Активно утворюються прикметники від іменників за допомогою префікса **поза-** і суфіксів **-н-**, **-ов-**, **-ськ-**. Такі слова використовуються на позначення економічних та політичних понять і явищ (*позаблоковий, позаєвропейський, позазаконний, позаринковий*).

У словотворенні нових прикметників велику роль відіграє префіксальний спосіб. СУМ-20 збільшився за рахунок активізації у процесах словотворення низки запозичених префіксів. Наприклад, дієвими елементами сучасної словотвірної системи є префікси: **анти-**, який використовується для творення переважно медичних (*антивірусний, антивітамінний, антимікробний, антиоксидантний*) та фінансово-економічних (*антидем-*

лінгвовий, антиінфляційний, антикомерційний) термінів; супер-, що вживається у лінгвістичній (суперлативний, суперсегментний) та технічній (супервізорний, суперелітний) термінологіях тощо.

Серед прикметникових інновацій у новому Словнику значну частину становлять складні форми прикметників. Високою активністю відзначаються при їх утворенні такі компоненти: науково- (науково-допоміжний, науково-експериментальний), теоретико- (теоретико-ймовірнісний, теоретико-концептуальний), селекційно- (селекційно-генетичний, селекційно-дослідний), семантико- (семантико-граматичний, семантико-морфологічний), санітарно- (санітарно-бактеріологічний, санітарно-гігієнічний, санітарно-епідеміологічний), електронно- (електронно-вакуумний, електронно-променевий), радіаційно- (радіаційно-екологічний, радіаційно-хімічний). При формуванні термінологічної лексики варто відзначити продуктивність числівника як першої частини складного прикметника (багатобітовий, двозарядний, однофазний, п'ятиелектродний, п'ятилопатевий, чотириатомний, шестициліндровий). З-поміж постпозитивних формантів активним виявився елемент -творчий (державотворчий, законотворчий, нормотворчий).

Порівняно з іменниками та прикметниками дієслівні новотвори у СУМ-20 демонструють меншу кількість одиниць.

Досить багато нових дієслів утворено за допомогою суфіксального способу. Найчисленнішу групу таких інноваційних одиниць становлять дієслова іменникового походження. При цьому простежуємо активність словотвірного типу з суфіксом -ува-. Новоутворена лексика представляє переважно наукову термінологію, наприклад: анестезувати (вет., мед.), алітувати (мет., хім.), вакуумувати (спец.), індукувати (фіз., ел.). Помітна також продуктивність суфіксів -изува-, -ізува- (амонізувати, бактеризувати, візуалізувати, максимізувати). Найбільшу і найрізноманітнішу групу відіменникових дієслівних утворень репрезентує загальнонаукова (догматизувати, параметризувати, сконкретизувати), економічна (комерціалізувати, приватизувати, тендеризувати), правова та політична (легітимізувати, майоризувати, політизувати) лексика. Ця група новотворів використовується також для нових найменувань у загальноживаній лексиці (актуалізувати, американізувати, корпоратизувати, урбанізувати).

Значну кількість дієслів становлять також одиниці, утворені за допомогою префіксів. У межах дієслівних неологізмів продуктивним є словотвірний тип із запозиченим компонентом ре-. Новоутворені лексеми переважно стосуються технічної сфери (реверсувати, рекуперувати, реструк-

турувати), а також фінансової та юридичної (*реінвестувати, рефінансувати*). Кількісно виділяється також словотвірний тип із префіксом *де-* (*дез-*): *дезактивувати, демонополізувати, дестабілізувати, десенсибілізувати*.

Серед дієслівних неологізмів, утворених на основі суфіксально-постфіксального способу, продуктивною виявилася модель із суфіксом *-ува-* і постфіксом *-ся* (*абсолютизуватися, бруталізуватися, політизуватися, структуруватися*). Утворені дієслова-інновації представляють переважно наукову та загальноживану лексику.

У СУМ-20 певну кількість нових слів становлять дієприкметники, похідні від дієслів недоконаного і доконаного видів. Вихідними компонентами для утворення дієприкметників цієї групи є дієслівні основи та суфікси *-н-*, *-ен-* (*-єн-*), *-т-*, (*вбавляти – вбавлений, вберегти – вбережений, вдавлювати – вдавлюваний, збільшити – збільшений, збризувати – збризнутий*).

Кількість нових прислівників у СУМ-20 порівняно з іншими частинами мови невелика. Це одиниці, що виникли переважно внаслідок префіксально-суфіксального способу творення. Такі прислівники – це передовсім слова, які існували у мові, але з різних причин не ввійшли до реєстру СУМ-11.

Більшість прислівників утворена від відносних та якісних прикметників із елементами *по-*, *-и* та *-ому*, де продуктивними є словотвірні типи із префіксом *по-* та суфіксом *-и* (*по-багацьки, по-крутійськи, по-хамськи, по-шкурницьки*) та з префіксом *по-* і суфіксом *-ому* (які мають також паралельні форми із суфіксом *-и*): (*по-діляцькому, по-екстремістському, по-ковбойському*).

У прислівниковій термінології переважає суфіксальний спосіб. Такі терміни здебільшого утворені від прикметників за допомогою суфікса *-о*: *алгоритмічно, апогамно, аспектологічно, білатерально, осмотично*.

Важливу складову частину СУМ-20 становлять реєстрові одиниці на позначення нових явищ, обумовлених зміною семантичної структури слів – семантичні неологізми. Можна виділити такі основні процеси семантичної неологізації: метафоризація, термінологізація та детермінологізація.

У межах семантичних явищ простежується, перш за все, активність метафоризації. Явище метафоризації суттєво торкнулося дієслівних (*випиратися, відлипати, відтягатися, навалюватися, обезкрилювати, позалптувати*) та іменникових (*амеба, архітектор, бритоголовий, відлига, гігант, індикатор, умілець*) номінацій, порівняно невеликою кількістю

представлені семантичні неологізми в групі прикметників (*ажурний, безпрограшний, борзий, класний, косметичний, тіньовий*).

Метафоричного переосмислення зазнав великий масив лексики, серед якої значне місце займають слова на позначення дій та процесів. Так, наприклад, увійшла в активне слововживання лексема *обламуватися*. СУМ-11 фіксує в ній таке значення: «утратити можливість продовжувати пересування, роботу через несправність машини, механізму і т. ін.». У процесі вживання цього слова в розмовній практиці розвивається нове значення «зазнавати невдач, розчарування у своїх планах, надіях і т. ін.» (*жарг.*). Словник демонструє також інші нові метафоричні утворення: *зліняти* «втративши попереднє забарвлення, набути невиразного, нерівного кольору; вицвісти». → «непомітно для всіх залишити місце роботи, ухилитися від якоїсь спільної діяльності» (*прост., жарг.*); *випиратися* «з силою пробиваючись, виходити назовні» → «ставати попереду інших» (*фам.*) та ін.

Велика кількість переносних значень з'явилася також в іменних найменуваннях, зокрема у словах на позначення осіб, у яких відображається ставлення до цих осіб суспільства й відповідна емоційна характеристика. Нові переносні значення слів – переважно з ремарками «розмовне» та «жаргонне», рідше «зневажливе» – постали в результаті їх уживання з відповідною оцінною характеристикою. Наприклад, слово *авторитет* «загально визнане значення, вплив, поважність (особи, організації, колективу, теорії і т. ін.)» та «особа, яка має вплив, повагу, заслуговує на повну довіру» набуло нового трактування: «лідер, ватажок кримінального угруповання» (*жарг.*); *умілець* «той, хто добре знає свою справу, володіє умінням; вправний майстер» → «спритна людина, що вміє вдало використовувати обставини» (*перен., розм.*) та ін.

Серед процесів семантичних змін актуальними виявилися також термінологізація, детермінологізація та ретермінологізація, які сприяли розширенню значення слів. При термінологізації відбувається перехід лексики із загальноживаного словникового складу до термінологічної сфери, при детермінологізації – із групи термінів до загальноживаних слів, а при ретермінологізації – перенесення терміна з однієї галузі в іншу з повним або частковим переосмисленням.

Завдяки розвитку комп'ютерної галузі виникло й закріпилося в мові нове значення слова *адмініструвати* «керувати установою, організацією, підприємством і т. ін.; управляти» → «здійснювати контроль та управління функціонуванням інформаційно-комп'ютерної системи, мережі» (*інформ.*). Значно оновилися також лексика технічної галузі, що свідчить про

розвиток цієї наукової сфери на сучасному етапі. Процеси формування термінологічних значень найпродуктивніші при творенні лексичних одиниць на позначення дій та процесів технічної галузі. Наприклад, у семантичній структурі слова *увивати* з'явилося нове значення «намотувати на когушку, барабан і т. ін.» (*техн.*). СУМ-20 репрезентує велику кількість семантичних неологізмів, які відображають розвиток лінгвістичної науки. Зокрема, значення слова *адресант* «той, хто адресує, надсилає кому-небудь листа, телеграму і т. ін., відправник» послужило основою для формування термінологічного значення «автор, учасник акту усної чи писемної комунікації» (*лінгв.*). Процес термінологізації слів відіграв помітну роль і у сфері медицини. Семантична структура слова *аплікація* «виготовлення орнаментів чи яких-небудь художніх зображень нашиванням або наклеюванням на що-небудь різнокольорових клаптиків тканини, паперу» розширилася на основі появи нового значення «накладання на хворий орган або місце певних речовин з лікувальною метою» (*мед.*).

Менш частотними виявилися семантичні неологізми, утворені шляхом детермінологізації і ретермінологізації. Приміром, слово *алерген* мало значення «речовина, що викликає алергію». Причиною його переосмислення стало широке використання зазначеної лексеми в мовленні, через що цей термін набув статусу загальноновживаного слова і, відповідно, нового значення – «те, що викликає draжливiсть; подразник» (*перен.*). Слово *абсорбція* розширило сферу функціонування шляхом використання в інших терміносистемах: *абсорбція* «поглинання, вбирання якої-небудь речовини із розчину або з газу всією масою іншої речовини» (*фіз., хім.*) → «вбирання тканинами молекул, які їх живлять» (*фізіол.*), «включення осіб, які прибули на постійне проживання в країну, в її економічне життя» (*екон.*).

Ще одна група інновацій у СУМ-20 – запозичення. Частина з них належить до загальноновживаної лексики: *брифінг, ваучер, грант, дилер, кілер, менеджер, ноу-хау, паркінг, плейс, провайдер, текіла, фритюр*. Термінологічних іншомовних слів найбільше припадає на медичну, технічну сфери, біологію, хімію, фізику. Наприклад, у медичній галузі запозиченнями представлені переважно назви захворювань, розладів і т. ін. людського організму (*абазія, варико́з, гемато́ма, гепатит, гінгівіт, глікемія, глікозу́рія (глюкозу́рія), грануле́ма (гранульо́ма), дисфазія, діарея, ектима, емпієма, енанте́ма, енурез, ілеус, імпети́го, ірит, орхіт, табес* тощо); у техніці – назви пристроїв, приладів, деталей та ін. (*віндроза, габiон, дефібра́тор, дефі́брер, думпер, екструдер, капсуль, сельсин, сканер, фітінг* тощо); у біології – речовин (*вісци́н, вітелiн, гіалiн, глобiн, еластiн, мiозин, пролан,*

пролін тощо), процесів, явищ (*гістоліз, гліколіз, мейоз*), клітин або складників їх будови (*диплоїд, оперон, стеригма* тощо) та багато інших.

Зазначимо, що активність неологізаційних процесів найменше позначилася на ядрі літературної мови – загальнонавчаних лексичі, яка виявила тенденцію до стабільності та стійкості складу. Свідченням цього є факт, що реєстр СУМ-20 поповнився тими загальнонавчаними лексемами (як правило, дериваційними утвореннями), які здебільшого існували в мові, але з тих чи інших причин не були зафіксовані в СУМ-11.

5.8. Стилiстична характеристика реєстрових слів

Стилiстичні ремарки у Словнику служать для характеристики тiєї частини складу сучасної лiтературної мови, яка обмежена у своєму вживаннi: використовується окремими групами людей, об'єднаними певною спiльнiстю, її функцiонування може бути обмежене певною ситуацiєю, сферою, часом i т. iн.

Ремарки подаються пiсля граматичної характеристики слова:

1) перед значенням слова, якщо воно одне:

ГАЛУШНИК... *жарт.* Любитель галушок...;

ІГНІТРОН... *фіз.* Іонний прилад, який має односторонню провідність і застосовується як потужний випрямляч змінного струму;

2) перед першим значенням багатозначного слова, якщо ремарка стосується слова в цілому:

БАЛАМКАТИ... *діал.* 1. Бити в дзвін; дзвонити... 2. Недбало махати...;

ЗЛАТИЙ... *заст., поет.* 1. Прикм. до злато... 2. *перен.* Щасливий...

3) перед значенням слова (після цифри), якщо ремарка стосується цього значення слова і всіх його відтінків:

БАЙДА... 1. *ч. і ж.* Безтурботна, гуляща людина; гульвіса... 2. *тільки ж., іст.* Довбаний човен... 3. *тільки ж., діал.* Шматок, грудка чого-небудь ламкого або крихкого (хліба, цукру, дерева)...

КАПАРИТИ... 1. *зневажл.* Робити нашвидку, абияк; партачити... // Погано куховарити. 2. Те саме, що *бідувати*...

Ремарки при усталених словосполученнях подаються після реєстрового фразеологізму перед тлумаченням, наприклад:

♦ **Впасти у лежку**, *рідко* – злягти, захворіти...;

♦ **Як (мов, ніби і т. ін.) зайцеві бубон**, *зі сл. потрібний, пром.* – уживається для повного заперечення змісту сказаного слова...

Набір стилістичних ремарок або стилістичних коментарів дає можливість характеризувати реєстрові мовні одиниці за функціонально-стильовими, часовими, емоційно-оцінними, частотними та іншими ознаками.

За допомогою функціонально-стильової характеристики відзначається закріпленість мовної одиниці за певним функціональним стилем, сферою вживання, соціальним середовищем.

Ремарка *розм.* характеризує одиниці, що належать до складу літературної мови, не виходять за рамки літературного слововживання, але використовуються переважно в усній мові та в мові персонажів художньої літератури при побутовому, повсякденному спілкуванні. Такі слова порівняно зі стилістично нейтральними одиницями, відзначаються певною невимушеністю, безпосередністю, «неофіційністю», наприклад:

БАГАТОТІРАЖКА... *розм.* Друкована газета, яку регулярно випускає великим тиражем яке-небудь підприємство, установа...

ВГ'АЧУВАТИ¹ (УГ'АЧУВАТИ)..., **ВГАТИТИ** (УГАТИТИ)... *розм.* Витрачати на щось у великій кількості (гроші, матеріал)...

ВАРКО... *розм.* Спекотно, душно...

Ремарка *книжн.* характеризує одиниці, що не властиві живій усній мові та використовуються звичайно у непобутових сферах спілкування (переважно в писемній мові), маючи стилістичні відтінки «вченості», піднесеності або архаїчності тощо. Наприклад:

АБРИС... *книжн.* 1. Обриси, контур... 2. *перен.* Загальна характеристика явищ, осіб, огляд подій і т. ін.

БЛАГОЛІПНИЙ... *книжн.* Гарний...

КОНТРОВЕРЗА... *книжн.* Суперечка, розбіжність думок, спірне питання; контроверсія...

Ремарка *поет.* вживається для позначення реєстрових одиниць, які використовуються у мові художньої літератури, переважно у поезії, замість загальноновживаних нейтральних, наприклад:

ВИНОГРОНО... *поет.* Гроно, кетяг винограду...

ЗАВЕСНІТИ... *поет.* Почати весніти...

СІЙНО... *поет.* Густо, часто...

Ремарка *нар.-поет.* вживається для позначення реєстрових одиниць, які зберігають своє фольклорне усно-поетичне забарвлення у складі літературної (звичайно поетичної) лексики, але не зливаються з нею, використовуючись як засіб стилізації, наприклад:

ДРІМЛІВИЦІ... *нар.-поет.* Те саме, що *дрімота*...

КРАСНО, *нар.-поет.* 1. Присл. до *красний*. ... 2. Щиро, сердечно, люб'язно...

КРУШИТИСЯ... нар.-поет. Журитися, тужити, страждати...

Позначки *біол., екон., лінгв., мед., спец., фіз.* та інші галузеві ремарки характеризують наукову термінологію та номенклатуру. Якщо термін уживається в двох або в трьох галузях, подається відповідна кількість ремарок в алфавітному порядку, наприклад:

АНТИСЕПТИЧНИЙ... *мед., фарм., хім.* Стос. до антисептиків....

ДЕПОЛЯРИЗАЦІЯ... *астр., фіз.* Усунення або послаблення гальванічної поляризації чи поляризації світла у світловому потоці....

ЕЛІПСОЇД... *астр., мат., фіз.* Замкнута овальна поверхня, утворена обертанням еліпса навколо однієї з його осей....

Якщо слово вживається більш ніж у трьох галузях, використовується ремарка *спец.*, наприклад:

БАРБОТУВАТИ... *спец.* Пропускати газ або пару крізь шар рідини для прогрівання або перемішування рідин....

ГІГРОСКОПІЧНИЙ... *спец.* Здатний легко вбирати в себе вологу....

Якщо характеристика галузевої належності слова міститься в його тлумаченні, відповідні стилістичні позначки звичайно не подаються, наприклад:

АБДОМІНОПЛАСТИКА... Хірургічна операція видалення надлишку шкіри та жиру живота (здебільшого у жінок)....

БАКЛАН... Водоплавний птах середньої величини родини бакланових звичайно чорного кольору з металічним відблиском....

Ремарка *діал.* характеризує одиниці, які, зберігаючи свою територіально обмежену належність, уживаються в мові художньої літератури або в усно-розмовній мові осіб, що спілкуються літературною мовою (головним чином для відтворення локального колориту, з певною зображально-виражальною метою або для позначення понять місцевого народного побуту, які не мають в літературній мові відповідника), наприклад:

БОКОРАШ... *діал.* Плотогон....

БОРШЕ... *діал.* Швидше....

БҀТНИЙ... *діал.* Пихатий, чванькуватий....

ВДВІЙЦІ (УДВІЙЦІ), ВДВІЙЗИ (УДВІЙЗИ)... *діал.* Вдвох....

Ремарка *зах.* характеризує одиниці, що вживалися в західноукраїнській літературно-мовній практиці протягом 2-ї пол. XIX – 1-ї пол. XX ст. які, не будучи вузькодialeкtnими, не ввійшли, однак, до загальноукраїнської літературної мови, наприклад:

АМБАСАДА... *зах.* Посольство, представництво....

ДУХОВИЙ... *зах.* Духовний (у 1 знач.)....

УЗГІДНОВАТИ..., УЗГІДНИТИ... *зах.* Погоджувати, узгоджувати....

Ремарка *прост.* вживається для позначення реєстрових одиниць, які, хоч і поширені на всій або значній території функціонування української мови, не входять, проте, до складу літературної мови (носіям літературної мови такі одиниці відомі, але звичайно відчуюються ними як ненормативні, субстандартні, знижені в своїй стилістичній характеристиці), наприклад:

НЕВАЖНІТЬКИЙ... *прост.* Досить неважний...;

ПОПУТКА... *прост.* Будь-який автомобіль, який їде в одному напрямку з ким-небудь, попутна машина...;

РОЗБУВАТИСЯ... **РОЗБУТИСЯ**... *прост.* Скидати із себе взуття...

Ремарка *жарг.* характеризує одиниці, які, вживаючись у середовищі певних соціальних, професійних або вікових груп мовців для номінації специфічних понять з криптологічною метою чи експресивним навантаженням, потрапляють до загальнонаціональної мови (до просторіччя, переважно міського, перебуваючи на межі з ним, до мови художньої літератури та публіцистики), зберігаючи, проте, в ній відповідний чужорідний статус. Наприклад:

ГЕНДЕЛИК... *жарг.* Невелике кафе, де продають спиртні напої...;

ЛАБУШНИЙ... *жарг.* Пов'язаний із музикуванням як заробітчанством...;

ШМОНАТИ... *жарг.* Обшукувати кого-небудь, порпатися в чий-небудь речах...

Часовою характеристикою супроводжуються мовні одиниці, що вже вийшли або виходять із живої сучасної мови. Така характеристика виражається як стилістичними позначками, так і коментарями в тлумачній частині статті.

Ремарка *заст.* характеризує назви предметів та явищ, що вийшли з ужитку носіїв мови або витіснені новими, сучасними назвами і сприймаються вже як архаїзми, наприклад:

АТРАМЕНТ... *заст.* Чорнило...;

БУРЛАКУВАТИ... I. *заст.* Не мати постійної роботи і постійного місця проживання, бути бурлакою (у I знач.)...;

ВАКАЦІЇ... *заст.* Канікули, перерва в роботі навчальних закладів, установ...;

ВЕСЬ²... *заст., уроч.* Село...

Емоційно-оцінні ремарки застосовуються для вираження емоційної оцінки – негативної чи позитивної.

Ремарка *зневажл.* використовується при мовних одиницях, що виражають відсутність поваги, нешанобливу, зверхню, образливу оцінку когось-небудь, наприклад:

БЛЮДОЛИЗ... *зневажл.* Той, хто догоджає кому-небудь, підлабузнюється заради власної вигоди; підлабузник...;

ЛАКЕЙСТВУВАТИ... *зневажл.* Бути лакеєм (у 2 знач.)...;

ПАРТАЦЬКИЙ... *зневажл.* Невміло або неохайно, недбало зроблений, виконаний...

Ремарка *лайл.* уживається для позначення реєстрових одиниць, які виражають різко негативне ставлення до об'єкта позначення або адресата в грубій і образливій формі, наприклад:

АРХАРІВЕЦЬ, АРХАРОВЕЦЬ... *лайл.* Бешкетник, босяк, волоцюга...;

КАРГА... *лайл.* Сварлива стара й потворна жінка...; відьма (у 2 знач.)...;

ПАСКУДНИЙ... 5. *перен., лайл.* Який викликає огиду...

Ремарка *ірон.* використовується для маркування реєстрових одиниць, що виражають тонку, приховану насмішку, зовні замасковану серйозною і позитивно-стверджувальною формою, прямо протилежною змісту, наприклад:

АВГУР... 2. *перен., ірон.* Людина, яка перетворює на тасмницю свої спеціальні знання; той, хто вдає, що він посвячений у надзвичайні тасмниці...;

ТАБАРИТ... 2. *перев. ми., перен., ірон.* Про комплекцію людини...;

КУКАТИ... 2. *перен., ірон.* Не знаходячи іншого виходу, перебувати, сидіти де-небудь, очікуючи кращих обставин, умов...

Ремарка *фам.* уживається при мовних одиницях, що виражають досить вільну, не зв'язану певними умовностями мовну поведінку, характерну для спілкування осіб, поєднаних близькими взаємостосунками, або зумовлювану бажанням створити таку обстановку при спілкуванні осіб, недостатньо близьких між собою, наприклад:

ВЕРЗТИ... *фам.* Говорити щось нерозумне, безглузде; вести пусті балачки...;

ГЕВАЛ... *фам.* Висока на зріст, неповоротка, незграбна людина...;

СИМПАТЯГА... *фам.* Симпатична, присмна людина...

Ремарка *жарт.* супроводжує мовні одиниці, що вживаються не всерйоз, а задля гумору, веселощів, розваг, не в образливій для когось-небудь формі, наприклад:

ГОЛОПУЗ... *жарт.* Про дитину...;

КАНДИБОБЕРОМ... *жарт.* Химерно, з викрутасами...;

МОКРОСТУПИ... *жарт.* Калоші...

Мовні одиниці з невеликою частотою вживання, з тих чи інших причин обмежені у своєму функціонуванні в рамках загальнонаціональної мови або певних її функціональних різновидів, характеризуються ремаркою *рідко* і *рідше*, наприклад:

АУТОДАФЕ, *рідше* **АВТОДАФЕ**... У середні віки – оголошення та виконання вироку інквізиції; звичайно публічне спалення еретиків і еретичних творів на вогнищі...;

ГОНОРУВАТИ, *рідко*. Триматися гордовито; зневажливо ставитися до інших...;

ГОРДЛИВИЙ... *рідко*. Те саме, що **горділливий**...

5.9. Текстово-ілюстративний матеріал

Для підтвердження функціонування слова в мові, наочнішого і повнішого розкриття його значення, синтаксичних зв'язків, уживання в певному словесному оточенні та з певним стилістичним забарвленням до реєстрових слів наводяться ілюстрації – приклади їх використання у літературній мові. При цьому ілюстративний матеріал не повторює зміст тлумачення реєстрової одиниці, а містить додаткову інформацію про неї.

У Словнику подано ілюстративні приклади двох типів: 1) цитати з художніх, наукових, публіцистичних та інших творів; 2) речення – словосполучення, які є типовими для вживання слова в літературній мові.

До багатозначних реєстрових слів та усталених словосполучень усіх типів ілюстрації наводяться, як правило, на кожне значення, виділене окремою цифрою чи буквеною позначкою. Проілюстровано також відтінки значень слів.

Ілюстрації наспоризуються короткою вказівкою на джерело (автор або тип публікації):

На полі вже дозрів бавовник білосніжний (Н. Забіла); Парламент Ісландії – альтинг – вважається одним із найдавніших у світі (з наук.-попул. літ.); Архалуки шилися приталеними, з глибоким вирізом на грудях, у чоловіків вони були довгополими, у жінок – короткополими (із журн.).

У новому Словнику витримано такий порядок розміщення джерел: з Біблії, з народної творчості (у тому числі з казки, з переказу), із збірок Номиса, зі словника Б. Грінченка, з художньої літератури (у хронологічному порядку за періодами творчості письменників), з Конституції України, з перекладної літератури, з мемуарної, публіцистичної, наукової, науково-технічної, науково-популярної, навчальної, релігійно-церковної літератури, з мови документів, із журналів, газет, з Інтернету, з мови реклами.

Після цього можуть подаватися речення у вигляді коротких вільних словосполучень.

До загальноновживаних слів та словосполучень цитати наводяться переважно з художньої літератури різних історичних періодів, починаючи з І. Котляревського й закінчуючи сучасними авторами (кінець XVIII – початок ХХІ ст.), подані у хронологічній послідовності відповідно до періоду творчості автора.

Для термінів і термінологічних словосполучень ілюстрації розміщені спочатку з наукової, науково-технічної, науково-популярної, навчальної літератури, а після них – із художньої літератури, періодики і т. ін.:

АСТМА, и, ж., мед. Задишка, спричинювана деякими захворюваннями серця або бронхів; ядуха. *Приступи серцевої астми часто настають уночі (з навч. літ.); Я здоров, хоч серце в мене в останній місяць таки .. позіршало, часті і болючі припадки астми (М. Коцюбинський)...*

Δ (1) Верблюжа колючка – пасовищна рослина роду напівчатарникових та багаторічних трав родини бобових, поширена у пустелях і напівпустелях Середньої Азії. *Коріння верблюжої колючки сягає ґрунтових вод на глибині до 16 м (з наук.-попул. літ.); Але немає на землі пустелі, в якій би не подавало ознак життя. У вигляді верблюжої колючки, павучка, вітру, що рухає бархани (А. Михайленко).*

Реєстрові слова можуть ілюструватися у складі прикладки:

АГА... *Жаба-ага, на відміну від інших земноводних, постійно зустрічається в солонуватих водах гирл рік, за що й отримала назву "морська жаба" (із журн.).*

В ілюстраціях до реєстрових прикметників та прислівників можуть подаватися відповідні прикметники та прислівники вищого і найвищого ступенів нормативного творення, наприклад:

КОРОТКИЙ... *Найкоротші дні бувають у червні (із журн.)...*

КРАСИВИЙ... *Хороші зірочки на небі, та вечірня зірочка красивіша над усім (Г. Квітка-Основ'яненко)...*

ГАРНО... *В хаті буде гарніше, як в церкві (Б. Харчук).*

У словникових статтях, які поєднують дієслівні видові пари, ілюстрації наводяться в межах кожного виду за прийнятою у СУМ-20 послідовністю слів у реєстровому ряді – спочатку на дієслово недоконаного виду як заголовне, а потім на дієслово доконаного виду:

ВИБИВАТИСЯ, аюся, аєшся, *недок.*, **ВІБИТИСЯ**, б'юся, б'єшся; *мн.* виб'ються; *док.* 1. З труднощами, докладаючи зусиль, вириватися, вибиратися звідки-небудь, уникати чогось. – *Пустя, Антосю! Годи-бо! Що ти? Побачать!* – *вибивається вона з обіймів (М. Старицький); Де сон староденний, мов*

крига, застиг, Творці вибиваються з темних кормиг (М. Рильський); Кидався [літак] вгору, обходив низом, намагаючись вибитися з-під прицілів зенітної артилерії (Ю. Яновський).

Те саме стосується і двовидових дієслів, у яких приклади вживання дієслова у недокопаному виді передують формам, що відповідають докопаному виду, наприклад:

АКУМУЛОВАТИ, *юю, юєш, недок. і док., що, спец.* Збирати, нагромаджувати, зосереджувати що-небудь. *Хлорофіл акумулює сонячну енергію (з наук.-попул. літ.); Завдяки спорудженню греблі Каховської ГЕС створено водоймище місткістю до 19 мільярдів кубометрів води, що дає змогу акумулювати наводкові води Дніпра і зберігати їх для корисної роботи (з наук.-попул. літ.); Сумні замріяні степи й похмуре місто, що акумулювало в собі гайдамаччину, не могли розрадити легковажної авантюристки – грекині (Б. Антоненко-Давидович).*

Хронологічний принцип подання ілюстрацій в межах кожного виду стосується і дієслівних фразеологізмів. Так, спочатку ілюструються в хронологічному порядку фразеологізми з дієслівним компонентом як вихідним у недокопаному виді, а після цього – з докопанним видом відповідного дієслова. При цьому порядок подання фразеологічних варіантів компонентів реєстрових дієслівних одиниць не враховується. Напр.:

◇ (2) **Вішатися (кідатися, чіплятися) / повіситися (кінутися, почепитися) на шию кому, несхв.** – нав'язуватися кому-небудь, настирливо домагатися прихильності, взаємності (про жінок у ставленні до чоловіків). – *Якби ти була моя жінка, а тут де не візьмись хвистик з уличці... І ти чіпляєшся йому на шию! Гарно було б мені на те дивитися? (Панас Мирний); – Ти, Яринко, йому на шию не вішайся, ми й почекаємо... У нас тих женихів хоч греблю гати... (М. Зарудний); – Сяя [Оксана] повісилась на шию капітанові [капітанові], захотіла разом панею бути (Г. Квітка-Основ'яненко); – Утік! Боягуз, утік від молодій жінки, що не почепилася йому на шию з першої зустрічі! (Іван Ле); Почувала [вона], що вільна вчинити, як хоче. Можє, й справді своє треба ловити, хапати? Інші самі ладні йому на шию кинутися, а вона ось таку владу над ним добула, доводить свого бригадира до безтяму... (О. Гончар).*

Але ж якщо у фразеологічній одиниці на першому місці як найуживаніший подано докопанний вид дієслова, то й, відповідно, наведені ілюстрації з фразеологізмами подаються саме цього виду. Напр.:

◇ (4) **Як (мов, ніби і т. ін.) із (з-під) землі вирости (вродитися, виринути) / виростати (виринати)** – несподівано, раптово з'явитися де-небудь. – *Як із землі виростає сестра Аркадія з пісним обличчям, з побожно згорненими на животі руками (М. Кошобинський); – Ох, лишенько! – задзвенів гурт. Ох, як цей козак злякав. – Де він взявся? – Наче з землі виріс (М. Лазорський); Враз біля нас як з-під землі виріс стражник Чапля (П. Панч); Скрізь по полю – ліво-*

руч і праворуч – стояли гармати, гармати, гармати... Немов уродилися з землі (О. Гончар); Славка.. наче з-під землі виринав (Л. Мартович); Як з-під землі виростає Стьопа (В. Собко);

Речення, які іноді наводяться для ілюстрації значень слів реєстру, подають, як правило, після ілюстрацій, зафіксованих у літературних джерелах:

АСКЕЗА, и, жс. 1. рел. Спосіб життя, що відповідає вимогам аскетизму (у 1 знач.). *Сповнене аскези і служіння Богу, життя Іоанна викликало велику повагу до нього жителів Іудеї, серед яких він проповідував (з наук.-попул. літ.); Монастирська аскеза;*

БЕЗВІДМОВНІСТЬ... Абстр. ім. до **безвідмовний**... *Тепер цінувалися тихість, безвідмовність, терплячість, тільки терплячим виконанням волі вищих можна було видертися з натовпу подібних собі (Ю. Мушкетик); Усе хлопчик приймав мовчки. Здається, що тая безодмовність дражнила хазяїна ще більш (Марко Вовчок); У монографії описуються найважливіші заходи забезпечення безвідмовності мікроелектронних схем (з наук. літ.); Підвищення безвідмовності металургійних машин.*

Деякі слова, переважно терміни, через відсутність ілюстративного матеріалу до них, не ілюструються:

АВІСТА, и, жс., фін. Напис на векселі або на якому-небудь іншому документі грошових розрахунків, який засвідчує, що вексель може бути оплачений на момент подання або через певний час від моменту його подання;

ДЖУБІЛІ, невідм., жс. Давня негритянська релігійна пісня-гімн американських негрів.

Якщо в ілюстрації потрібні уточнення, то використовуються квадратні дужки, в яких подано слова, що розкривають значення:

– відсутніх у цитаті підмета й додатка, якщо це породжує неясність, двозначність:

У кого є голос могутній, дзвінкий, Нехай на весь світ про Вітчизну співає, І співом вабливим докупі еднає [народ] (Дніпрова Чайка); Через два літа стояв [Святополк] під Києвом з печенігами, які йшли до нього, мов собака на пошвист, ваблені обіцяними грабунками багатого стольного города (П. Загребельний); Один чи два рази проказала [Настя] за Чумаком абетку і вже на другий день, дивуючись сама з себе, назвала всі літери (В. Речмедін).

– займенників (він, вона, воно, вони, сам, сама та ін.), займенникових прислівників (там, туди, тоді та ін.) для розуміння контексту:

І, взявши його [апостола Павла], повели в ареопіг та й казали: Чи можемо знати, що то є ця наука нова, яку проповідуєш ти? (Біблія. Пер. І. Огієнка); Лев бовтався там [у колодязі], поки й захлинувся (з казки).

У всіх інших випадках до подібних займенників та прислівників уточнення не подаються:

БЕЗВИХІДЬ... Безвихідне становище. *Він у безвиході весь час вірив, що то був знак долі, яка вихопить його із содому страху* (Р. Іваничук)...

Якщо в реченні використано лексеми, які не відповідають нормам сучасної літературної мови та чинному правопису, то в дужках зазначено їх нормативний відповідник:

До того люд домуштровали [домуштрували]... (Т. Шевченко); *І усміхнувшись якимсь безвиразним усміхом, дивилась я на неї широко отвертими [відкритими] очима* (О. Кобилянська).

Для зменшення обсягу словникової статті цитати з художньої літератури можуть скорочуватися. Пропуски позначаються двома горизонтальними крапками (...), які наводяться лише всередині ілюстрації:

Антон ... стояв у полотняній сорочці, старанно вилатаній на плечах (С. Чорнобривець); *Брати у батьківській жили старій оселі, .. Але з-між них один Усе, було, пісні вигадає веселі* (М. Рильський).

Якщо скорочується початок чи кінець ілюстрації, такі пропуски не відзначаються.

Якщо в джерелі, з якого взято ілюстрацію, речення закінчується трьома крапками, то вони залишаються і в ілюстрації, наприклад:

Заснув Чабан, а лишенько не спало: Гадюка вилізла з трави, сичить... (Л. Гіттов); *— Вони шпигають, що в мене голуба кров та руки аристократа і я не люблю простого народу...* (В. Дрозд).

Порівняно з СУМ-11 джерельна база СУМ-20 значно збільшилася. Зокрема, додано ілюстрації з творів майже 150 письменників новітньої доби, а також тих, чий доробок не внесено до 11-томника з ідеологічних міркувань. У повному списку джерел нового Словника, вміщеному в Додатку, прізвища цих письменників виділено курсивом.

Новим інструментом для добору ілюстрацій словникових статей СУМ-20, який не існував під час укладання СУМ-11, є Український національний лінгвістичний корпус. Його будову, функціональні особливості та застосування описано в підрозділі 6.4.

5.10. Поліграфічне оформлення СУМ-20

Поліграфічне оформлення будь-якого словника розв'язує два основних завдання. З одного боку, воно спрямоване на втілення принципів видавничої естетики, що робить спілкування зі словником для користувача комфортним. Другим і основним завданням поліграфічного оформлення є

однотипне позначення однорідних елементів структури словника, яке полегшує користувачеві роботу з ним, орієнтацію в словниковій інформації, її знаходження. Академічні словники, як правило, характеризуються науково опрацьованою, розгалуженою словниковою структурою, у тому числі структурою словникових статей, у якій втілено принципи певної лексикографічної теорії. Засоби поліграфічного оформлення тут виступають ідентифікаторами окремих елементів такої структури, формуючи так звану метамову словника, тобто формальну мову опису структури його лексикографічної системи.

У такий спосіб елементи словникових статей набувають свого формального визначення, що дозволяє за наявності електронного тексту словника скомпонувати в автоматичному режимі відповідну лексикографічну базу даних.

Ось чому при укладанні СУМ-20, прототипом якого був СУМ-11, приділено таку велику увагу описові поліграфічного оформлення та інтерпретації його елементів як елементів метамови СУМа.

Шрифтові виділення

I. Напівжирним прямим шрифтом набрані:

1. Великими літерами:

а) реєстрові слова (**БАРАБАН, ЗМІ, АВТО...**);

б) фонетичні варіанти реєстрового слова, розроблені в одній словниковій статті: **АКА́НТ, АКА́НФ; ВНИЗ (УНИЗ)**. Зауважимо, що фонетичні варіанти реєстрових слів подаються в одній статті через кому, утворюючи реєстровий ряд, а позиційні фонетичні варіанти оформляються у круглих дужках;

в) видові пари реєстрових дієслів або ті дієслова, що різняться суфіксами чи наголосом, розроблені в одній словниковій статті: **ВИХО́ДИТИ..., ВЙ́ЙТИ; ПІДКУ́ПОВУВАТИ..., ПІДКУПА́ТИ..., ПІДКУПЛЯ́ТИ..., ПІДКУПІ́ТИ; ВИСМИКА́ТИ..., ВИСМИ́КУВАТИ..., ВИСМИКА́ТИ..., ВИСМИ́КНУТИ.**

2. З великої літери:

а) перші букви стійких словосполучень (**Віде́нський вальс**);

б) словосполучень – еквівалентів слова (**○ До безме́жжя**);

в) термінологізованих (**△ А́вторський а́ркуш**);

г) фразеологічних словосполучень (**◇ Вто́піти в ло́бці води**).

3. Малими літерами:

а) усі типи словосполучень (крім першої літери) з можливими варіантами: **◇ Воді́ти / поводі́ти за ніс (за но́са, рідко за кі́рпу)...**

б) форми однини в іменникових статтях з реєстровим словом у формі множини: **БЕЛЬГІЙЦІ**, ів, мн. (одн. **бельгієць**, йця, ч.; **бельгійка**, и, жс.)...; або, навпаки, в однині (у деяких випадках): **ОКО**¹, а, с. (мн. **оці**, рідко **вічі**...)...

в) форми вищого і найвищого ступенів прикметників та прислівників, якщо вони уведені в статтю: **ГАРНИЙ**... *вищ. ст. кращий і ліпший, рідко гарніший...*

г) реєстрові слова, на які даються посилання у відсильних тлумаченнях різного типу:

- Те саме, що...: **АВТОСТРАДА**, и, жс. Те саме, що **автомагістраль**...
- Прикм. до...: **ГУЦУЛЬСЬКИЙ**, а, е. Прикм. до **гуцулі**...
- Дія за знач.: **КАЛІБРУВАННЯ**, я, с. *спец.* Дія за знач. **калібрувати**...
- Зменш.-пестл. до...: **БИЛИНОЧКА**, и, жс. Зменш.-пестл. до **билінка**...

тощо;

г) слова у відсильних статтях після див.: **УКРАЇНЕЦЬ** див. **українці**;

д) слова, які входять до синтаксичних конструкцій на зразок: у знач. ім. (у знач. ім. **поранений**), у знач. вставн. сл. (у знач. вставн. сл. **головне**), у знач. присл. (у знач. присл. **амфітеатром**) тощо;

е) безособові предикати, розроблені як відтинки до відповідних дієприкметників (// **вбито**, *безос. пред.*).

4. Маркери омонімії – нарядковими арабськими цифрами: **КОСА**¹, **БАБКА**¹⁰.

II. Світлим прямим шрифтом набрані:

1. Граматичні форми, подані при реєстрових словах: **ПІСТ**, **посту**...;

БІЛИЙ, а, е...; **ВИТИ**¹, **вію**, **вієш**...

2. Тлумачення: **ФІЗИК**... Фахівець із фізики...

3. Номери значень полісемічного слова, на яке є відсилання у відсильних тлумаченнях: **ГУРТКОВИЙ**... Прикм. до **гурток** 2...

4. Буквене позначення у полісемічних словосполученнях:

○ **Був та загів (поплів)**: а) зовсім зник, пропав, не стало кого-, чого-небудь...; б) перестав жити, існувати...; в) не мав можливості здійснитися (про задум, мрію і т. ін.)...

5. У квадратних дужках – уточнення в цитатах: **Заклінав** [Демид]; нормативні вживання при позанормативних словах: у його [нього]; а також пояснення при словах, яких немає в реєстрі: **кливаком** [іклом].

6. У круглих дужках – джерела ілюстрацій: (Т. Шевченко), (з наук. літ.), (з газ.)

7. Сполучник «і», якщо він виступає між двома напівжирними твірними словами, вжитими у тлумаченні:

ОРІЕНТУВАННЯ... Дія за знач. **орієнтувати** і **орієнтуватися...**;
КОНОПЛЯРСЬКИЙ... Прикм. до **конопляр** і **коноплярство...**

8. Функціональні позначки до ілюстрацій (* Образно, * У порівн.):
* Образно. *Бідою знов небо дихало*; * У порівн. *Парубки ж у нас, як орлята: жваві, молоді.*

III. Світлим курсивом набрані:

1. Позначки граматичних форм і сполучуваності слів, наприклад: *ч., ж., мн., док., недок., невідм., безос., у знач. вставн. сл., кого, з інфін., зі сл., з ким – чим, на що, по чому.*
2. Ремарки, наприклад: *розм., книжн., спец., рідко, вет., мед., жарг.*
3. Відсилання див.
4. Сполучник «і» між деякими позначками граматичних форм і сполучуваності слів:

ПОАГІТУВАТИ, *ую, уєш, док., кого, що і без прям. дод. ...*; **СІМ**, *семі і сімох...*, **ВІХОР**¹, *хору і хру...*

5. Скорочення і т. ін. в усталених словосполученнях: **Брати** (*забирати, перехоплювати і т. ін.*) / **взяти** (*забрати, перехопити і т. ін.*) **ініціативу...**; при переліку парадигматичних форм: **ВОНІ**, *їх (з прийм. до них, від них і т. ін.)...*

6. Ілюстрації, в тому числі речення:

Перші українські словники кінця XVI – початку XVII ст. мали значення своєрідних енциклопедичних довідників (із журн.); Абдуктивний метод.

7. Авторські ремарки у драматичних творах, які є ілюстраціями до реєстрових слів і фразеологізмів: [Груїчева (*складає карти*):].

IV. Урозбивку набрано:

1. Пояснення до деяких тлумачень на зразок «за знач. близький до але».

2. Приклади до реєстрових слів, що становлять першу або останню частину складних слів із продуктивними словотвірними компонентами: **СВІТЛО**... Перша частина складних слів... *світловідбивач, світловідчуття, світлодайний...*

3. Слова, які вказують на синтаксичну сполучуваність реєстрових лексем або фразеологізмів:

ЗАХОПЛЮВАТИ... *2. також зі сл. з собою...*

Як (*мов, ніби і т. ін.*) **біда на голову, зі сл. з'явиться, звалитися і т. ін.**...

4. Назви функціонально-семантичних відношень у деяких применникових статтях:

ЗА¹, прийм.; з род., знах. і оруд. в. Сполучення з прийм. за виражають: Просторові відношення...

5. Антонімі у тлумаченнях, які наведено після слова «протилежне»: **СКОРБОТА**... печаль, сум; протилежне радість...

6. Назви дійових осіб в ілюстраціях із драматичних творів:

[Зінька:] *Ти, може, не захочеш і похристосуватися зо мною?* [Роман:] *Чому? Це ж закон... Давай похристосуємось* (М. Кропивницький).

Використання дужок

I. У круглих дужках подано:

1. Фонетичні варіанти реєстрових слів: **ІТИ (ЙТИ), ВСТРЯГАТИ (УСТРЯГАТИ)**.

2. Форми однини з відповідним граматичним оформленням у статтях із реєстровим словом у формі множини: **БЛОРУСИ**, ів, мн. (одн. **білорус**, а, ч.; **білоруска**, и, ж.)...; або множини, де реєстрове слово розроблене в однині: **ОКО¹**, а, с. (мн. **очі**, рідко **вічі**, очей – перев. із прийм. у)...

3. Вказівка на номер значення полісемічного слова, вжитого у тлумаченні, якщо це слово набрано світлим шрифтом: **ГУРТКІВЕЦЬ**... Член гуртка (у 2 знач.)...; на омонімічне слово, через яке тлумачиться відповідне реєстрове слово: **АВЕРСНИЙ¹**... Стос. до аверсу (див. **áверс¹**)...; в усталених словосполученнях із тлумаченнями на зразок «Те саме, що...» – для зазначення слова, де вони розроблені: \diamond **Позабувати** язикі *десь* – те саме, що **Забутьи** язикі *в роті* (про всіх або багатьох) (див. **забувати**)...

4. Парадигматичні форми при вживанні реєстрового слова у певній синтаксичній функції, наприклад: **ЗДАВАТИСЯ²**... 3. *тільки здається* (здавалося), у знач. вставн. сл. ...

5. Уточнення на зразок (з великої літери): **БОГ**... 1. род. Бóга, бібл. (з великої літери)...

6. У всіх типах стійких словосполучень – компоненти, крім факультативних: \diamond **Здіймати / здійснити (зняти) [важкий] кáмінь (тягáр) з сéрця (з душі)**.

7. Вказівки на синтаксичну сполучуваність при багатозначних усталених словосполученнях:

Брати / взяти на замітку: а) (кого і без дод.) брати кого-небудь під особливий нагляд...; б) (що і без дод.) враховувати що-небудь...

8. Порядкові номери стійких словосполучень у межах статті: \diamond **Багато (забагато) чéстí**...

9. Джерела ілюстрацій: (Леся Українка), (із журн.), (з Інтернету).

II. У квадратних дужках подано:

1. Вставки в цитати для уточнення: *він* [Іван], *тоді* [торік] або в разі ненормативного слововживання: *ходе* [ходить], *з радіщ* [радощів].

2. Імена або назви дійових осіб в ілюстраціях із драматичних творів:

[Мавка:] *Я – Мавка лісова.* [Лукаш:] *А, от ти хто! Я від старих людей про мавок чув не раз* (Леся Українка).

3. Факультативні компоненти сталих словосполучень:

◇ Як (мов, ніби і т. ін.) [та] ба́ба Пала́жка [і ба́ба Параска]...

III. У кутових дужках подано варіанти переважно фразеологічних словосполучень, у яких варіант одного з компонентів приводить до граматичної зміни іншого: Як [водá] з прорвано́ї грéблi <Як [водá] крізь прорвану грéблю>...; Сумні́в берé / узя́в <Сумні́ви беру́ть / взяли>...; а також усталені словосполучення з інверсійним порядком компонентів: ◇ І на́шим і ва́шим <І ва́шим і на́шим>...

Знаки пунктуації як поліграфічні маркери

I. Крапка (.) ставиться:

1. У кінці лівої частини статті перед тлумаченням: **ВІСК**, во́ску, ч. 1. Органічна речовина...

2. Після цифри рубрики: **ПІСНЯ**... 1. Словесно-музичний твір, признач. для співу...

3. Після тлумачення: **ВІШАНИЙ**, а, е. Дієпр. пас. до **вішати**. *Постійно битий, різаний, вішаний, голодом морений народ...*, крім випадку відсутності ілюстрації перед наступним відтінком: **АКТУАЛІЗОВАНИЙ**... Дієпр. пас. до **актуалізува́ти**; // у знач. прикм. ...

4. Після позначок *Образно. і *У порівн.

5. У кінці абзацу.

6. Після скорочень: *напр., прикм., спец., екон.* і т. ін.

7. У кінці словникової статті.

II. Кома (,) ставиться:

1. Перед граматичними формами реєстрових слів: **БІСКВІТ**, а, ч. ...

2. Між кількома граматичними формами: **ВЕРСТАТИ**, аю, аєш, *недок., що...*

3. Між заголовним словом та його варіантами: **АНАЛГЕТИКИ**, **АНАЛЬГЕТИКИ**, ів, мн. (*одн. аналгетик, анальгетик, у, ч.*), *фарм. ...*; **НАЙМАТИ**, аю, аєш, *недок.*, **НАЙНЯТИ**, найму́, наймеш, *док. ...*

4. Перед стилістичними ремарками і граматичними позначками: **БАЙТ**, а, ч., *інформ.* ..., крім вказівки на синтаксичну сполучуваність в усталених словосполученнях: \diamond **Завершальний (останній) акорд чого і без дод.**

5. Між кількома ремарками: **ВАКЦИНУВАТИ**, *ую, усш, недок. і док., ко-го, вет., мед.* ...

III. Крапка з комою (;) ставиться:

1. У дієслівній статті перед позначками на зразок: *мн., мин. ч., наказ. сп.:*

БРАТИ, *беру, береш; недок.* ...

2. Між наведеними у дужках формами однини чоловічого і жіночого роду іменників, розроблених при множинному іменникові, з відповідними відмінковими закінченнями та позначеннями роду: **АБІССІЙЦІ**, *ів, мн. (одн. абіссієць, пця, ч.; абіссійка, и, ж.)...*; між поданими як відтінки субстантивованими прикметниками у різних родах, а також у множині, розробленими при реєстровому прикметникові або дієприкметнику: ... // у *знач. ім. хворий, рога, ч.; хвора, рої, ж.; хворе, рога, с.; мн. хворі, рих...*

3. У тлумаченнях:

а) перед частиною, яка доповнює описове тлумачення: **АБЕТКА**...

2. Книжка для початкового навчання грамоти; буквар...; **АГАТ**... Мінерал із шаровим розподіленням забарвлення...; різновид халцедону...

б) перед словом «протилежне», яке передує антонімам: **АБСТРАКТНИЙ**... 1. Який виник у результаті абстрагування; протилежне конкретний...

4. Між ілюстраціями.

5. Між усталеними словосполученнями всіх типів.

6. Перед відтінками значень.

7. Перед * **Образно**, * **У порівн.**, якщо перед ними є ілюстрація.

8. Між підрубриками а), б) і т. ін. в усталених полісемічних словосполученнях усіх типів та у прийменникових статтях.

9. Між підрубриками 1), 2) і т. ін. у статтях із продуктивним словотвірним компонентом: **АВТО**...¹ Перша частина складних слів, що відповідає: 1) слову **автомобільний**, наприклад: автобаза, автовокзал...; 2) слову **автомобіль**, наприклад: автобензовоз, автоцистерна.

IV. Двокрапка (;) ставиться:

1. Після реєстрового слова (з можливими граматичними формами і позначками), яке вживається лише у складі фразеологізму: **ВАРФОЛОМІВСЬКИЙ**, а, е: \diamond **Варфоломівська ніч** *див. ніч.*

2. Перед підрубриками а), б) і т. ін. у всіх полісемічних усталених словосполученнях:

◇ Впа́дати (рідше па́дати) / впа́сти в бо́ко (в бо́чі, у ві́чі): а) (кому) хто-небудь помічає, бачить, звертає увагу на когось, на щось...; б) (кому) привертати чию-небудь увагу...

3. Перед підрубриками 1), 2) і т. ін., а також перед прикладами у статтях із продуктивним словотвірним компонентом: АВІА... Перша частина складних слів, що відповідає: 1) слову авіаційний, наприклад: авіабензін...;

4. У ремарках в драматичному творі: [2-ий парубок:], [Марія (встає):].

V. Тире (–) ставиться:

1. При вказівці на синтаксичну сполучуваність, коли пропущений прийменник *за кого – що, перед ким – чим*.

2. Перед тлумачною частиною однозначних усталених словосполучень усіх типів: А¹... ◇ (1) Від а до я (до зет) – від початку до кінця; все повністю...

3. Замість лапок перед прямою мовою: АЗБУКА... – Дисципліни без авторитету нема – невже цієї азбуки вас не вчили? (О. Гончар)...

VI. Дві крапки (..), які відокремлюються від сусідніх слів пробілами, ставляться всередині ілюстрацій на позначення їх скорочення із зазначенням прізвища автора. Якщо ілюстрація переривається після розділового знака, то пробіл ставиться після нього, якщо перед розділовим знаком – пробіл перед ним не ставиться: *Єремія, вихований ще в єзуїтських школах... прийнявши латинство, став .. гонителем православного люду* (О. Стороженко).

Шрифт знака, який пишеться без пробілу після слова, збігається зі шрифтом попереднього символу.

VII. Скісна лінія (/) ставиться у дієслівних словосполученнях для відокремлення видових пар: ... Давати / да́ти віддіт...

VIII. Дві скісні лінії (//) ставляться для позначення відтінку значення: ВЕЖА... 1. Висока вузька споруда...; // Вишка для спостереження або передачі сигналів...

З абзацу починається:

1. Словникова стаття в цілому.

2. Кожна рубрика статті, крім першої.

3. Кожен тип усталених словосполучень.

Особливості поліграфічного оформлення покажчика словосполучень подано в наступному підрозділі.

5.11. Індекс СУМ-20

На відміну від традиційних видань багатотомних тлумачних словників кожний друкований том СУМ-20 супроводжується блоком інформаційно-пошукових матеріалів, до яких належать список реєстрових слів (СРС) і покажчик словосполучень (ПС). Ці, певним чином структуровані, тексти виконують роль індексу словникової системи – вони організують додатковий доступ до структурних елементів словникових статей.

СРС конкретного тому містить усі слова, наявні в його реєстрових рядах. Нижче подані приклади реєстрових (лівих) частин словникових статей, які мають більш ніж одне слово в реєстровому ряді:

АїР, *діал.* **АєР**, у, ч.

...**АКТНИЙ**, *рідко* ...**АКТОВИЙ**.

БАБАХАТИ, **БАБАХКАТИ**, аю, аєш, *недок., розм.*

ЖОДЕН, *дна, дне*, **ЖОДНИЙ**, а, е, *заст.* **ЖАДЕН**, *дна, дне*, **ЖАДНИЙ**, а, е, *займ., запереч.*

УРОЖАЙ (**ВРОЖАЙ**), аю, ч.

Окрему групу становлять словникові статті з реєстровим іменником у множині на позначення назви народу, народності, етнічної групи і т. ін., а суб'єкти такої єдності, представлені іменниками в чоловічому та жіночому роді однини або лише в чоловічому, подаються в дужках з відповідними ремарками:

АВАРЦІ, *ів, мн. (одн. аварець, рця, ч.; аварка, и, ж.)*.

АГУЛІ, *ів, мн. (одн. агул, а і агулець, лья, ч.; агулка, и, ж.)*.

ВАРЯГИ, *ів, мн. (одн. варяг, а, ч.)*.

Одиниці реєстрових рядів супроводжуються відсиланнями на заголовне слово, тобто перше слово реєстрового ряду. Заголовне ж слово до одиниць реєстрового ряду не відсилається. Реєстрові слова у списку подаються великими літерами з наголошенням та маркером омонімії, заголовне слово у відсиланні наводиться малими літерами. ПС надає користувачеві можливість знайти словосполучення за будь-яким його повнозначним компонентом. Нижче подано фрагмент СРС, що відповідає прикладам реєстрових рядів:

АВАРЕЦЬ *див. аварці.*

АВАРКА *див. аварці.*

АВАРЦІ

АГУЛ *див. агулі.*

АГУЛЕЦЬ *див. агулі.*

АГУ́ЛИ...

АГУ́ЛКА *див. агу́ли.*

АДАМІ́Т *див. адамі́ти.*

АДАМІ́ТИ...

А́ІР *див. а́ір.*

А́ІР...

АКА́НТОВИЙ...

АКА́НФОВИЙ *див. ака́нтовий.*

... А́КТНИЙ...

... А́КТОВИЙ *див. ...а́ктний.*

АМБІ...

АМБО... *див. амбі.*

АНА́ФЕМА...

АНА́ХТЕМА *див. ана́фема.*

АПРОБÓВУВАТИ...

АПРОБУВА́ТИ *див. апробóвувати.*

АРЕШТÓВУВАТИ...

АРЕШТУВА́ТИ *див. арештóвувати.*

Місце розташування реєстрового словосполучення у тексті друкованого словника визначається двома параметрами: заголовним словом словникової статті та порядковим номером словосполучення. Сукупність цих параметрів є адресою реєстрового словосполучення (АРС).

Заголовне слово в АРС подано в його повній реєстровій формі – із наголошенням та позначкою омонімії. Номер словосполучення наведено в тій самій формі, що у фразеологічному блоці — в круглих дужках після заголовного слова.

Для кожного реєстрового словосполучення будується елемент покажчика із такою структурою: ідентифікація типу словосполучення, подання словосполучення у його повній словниковій формі, службове слово *див.* та АРС. Наприклад:

Бра́ти / взя́ти ако́рд *див. бра́ти (1)*

△ Адре́са електро́нної по́шти <Електро́нна адре́са> *див. адре́са (2)*

○ До безме́жжя *див. безме́жжя (1)*

◇ Від а́за до і́жиці *див. аз¹ (1)*

ПС кожного тому формується зі словосполучень, повнозначні компоненти яких належать до реєстрового ряду словникових статей цього тому. У відсильних словосполученнях порядковий номер не ставиться, наприклад: В абе́ткóвому (алфа́вітному) поря́дку *див. поря́док.*

Словосполучення, до складу яких уходить заголовне слово в будь-якій словозмінній формі, утворюють гніздо. Заголовне слово гнізда – це пов-

нозначний компонент словосполучення у вихідній граматичній формі: іменники подано в називному відмінку однини (крім множинних іменників); для прикметників, дієприкметників та займенників вихідна форма — називний відмінок однини чоловічого роду; для дієслів — інфінітив. Кожне словосполучення подається у покажчику стільки разів, скільки має повнозначних компонентів. Усередині гнізда словосполучення розподіляються за відмінковими формами. Наприклад, термінологічне словосполучення Δ Адреса електронної пошти <Електронна адреса> входить до гнізд із заголовними словами АДРЕСА, ЕЛЕКТРОННИЙ та ПОШТА:

АДРЕСА...

адреса

Δ Адреса електронної пошти <Електронна адреса> *див.* адреса

ЕЛЕКТРОННИЙ...

електронна

Δ Адреса електронної пошти <Електронна адреса> *див.* адреса

електронної

Δ Адреса електронної пошти <Електронна адреса> *див.* адреса

ПОШТА

пошта

Δ Адреса електронної пошти <Електронна адреса> *див.* адреса

Усі структурні компоненти покажчика — заголовні слова гнізд, відмінкові форми заголовних слів та самі словосполучення упорядковуються на відповідному ієрархічному рівні за алфавітом. При побудові ПС автоматизуються такі процеси: нумерація розроблених словосполучень, формування елементів покажчика та групування їх у гнізда, розподіл словосполучень за словозмінними формами заголовного слова, упорядкування за алфавітом. Розподіл за словозмінними формами здійснюється за допомогою електронного словника словозміни української мови. Наприклад:

АВТОРСЬКИЙ...

авторська

Δ Авторська школа *див.* школа

авторське

Δ Авторське право *див.* право

Δ Авторське свідоцтво *див.* свідоцтво

авторський

Δ Авторський аркуш *див.* аркуш

Δ Авторський примірник *див.* примірник

Δ Авторський (творчий) концерт (вечір) *див.* концерт

Показчик дає можливість працювати зі словосполученнями як із незалежною словниковою підсистемою. Він виконує такі функції: 1) полегшує користувачеві пошук місця лексикографічного опрацювання усталеного словосполучення у Словнику; 2) забезпечує зв'язок між фразеологічними підсистемами окремих томів; 3) як окремий структурований текст показчик дає змогу простежити закономірності представлення лексики в реєстрових словосполученнях.

6.1. Парсинг¹ СУМ-11

Як зазначено вище, для укладання нового 20-томного тлумачного Словника української мови в Українському мовно-інформаційному фонді НАН України розроблено нову – комп'ютерну – технологію укладання тлумачних словників. Її реалізовано у вигляді так званої фундаментальної академічної лексикографічної системи «Словник української мови» (СУМ), яка містить інструментальний комплекс, орієнтований на комп'ютерне розв'язання завдань тлумачної лексикографії. У складі зазначеного комплексу міститься добре структурована лексикографічна база даних (ЛБД) тлумачного словника та реалізовано низку лексикографічних і технологічних функцій.

Робота зі створення інструментального комплексу «СУМ» проводилася в кілька етапів. Перший етап передбачав переведення тексту СУМ-11 до цифрової форми і виконувався засобами сканування та розпізнавання тексту, в результаті чого одержано електронний варіант 11-томного словника. Ланцюг підготовки СУМ-11 до конвертації в ЛБД представлено на рис. 6.1.

Після сканування і розпізнавання тексту СУМ-11 (9856 сторінок) його збережено в RTF-форматі та роздруковано з метою виправлення помилок, які виникли під час роботи програми оптичного розпізнавання.

Потім здійснено подвійну коректуру і внесено виправлення до електронного тексту СУМ-11, внаслідок чого отримано його еталонний варіант.

¹ Парсингом словника називають автоматичний (автоматизований) процес переведення його тексту до комп'ютерної лексикографічної бази даних із певною структурою.

Останній етап підготовки до створення ЛБД – конверсія з RTF- до HTML-формату із системою кодування Unicode засобами текстового редактора MS Word.

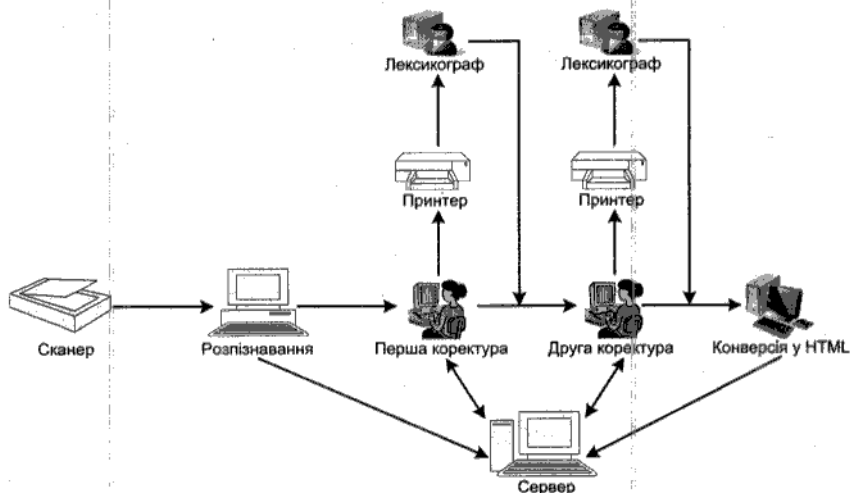


Рис. 6.1. Ланцюг підготовки тексту СУМ-11 до конвертації в ЛБД

Слід відзначити, що робота з ЛБД у форматі електронного тексту такого значного обсягу (понад 135 мегабайтів дискового простору) є абсолютно неефективною. Окрім дуже повільного темпу, з яким система працює на великих текстових масивах, принципову незручність становить неможливість прямого доступу до окремих елементів структури Словника. Такі вади можуть звести нанівець усі переваги, які надає використання комп'ютерів у лексикографії. Отже, постала необхідність розробки спеціалізованої ЛБД Словника української мови, причому її формування треба було здійснити автоматично, оскільки в ручному або напівручному режимі створення такої бази даних, виходячи з електронного тексту, як переко-нує досвід, є неможливим.

Для забезпечення процесу автоматичної конверсії тексту СУМа до ЛБД – а саме в цьому й полягає зміст парсингу Словника – в Українському мовно-інформаційному фонді НАН України розроблено програмне забезпечення виділення елементів його структури відповідно до будови лексикографічної системи СУМа з використанням поліграфічних ознак її текстової ідентифікації, наведених у табл. 6.1.

Таблиця 6.1. Поліграфічні ознаки ідентифікації елементів структури СУМа

№	Структурний елемент	Формальна ознака структурного елемента
1	Реєстрове слово N	Початок: абзац, літера (українська, напівжирний шрифт, верхній регістр). Закінчення: кома; крапка; пробіл, індекс, цифра, кома; пробіл «див.»; пробіл, індекс, цифра, пробіл «див.»; двокрапка.
2	Ліва частина статті N, початок	Реєстрове слово N
3	Частини мови:	<i>Ознака частини мови у лівій частині</i>
	Іменник	<i>с.;</i> <i>ж.;</i> <i>ч.;</i> <i>мн.</i> При аналізі <i>с.</i> перевірка на наступні пробіл та <i>ч.</i>
	Прикметник	<i>прикм.;</i> <i>вищ. ст.;</i> <i>збільш.;</i> <i>змени.;</i> <i>змени. -пестл.</i>
	Числівник	<i>числ.</i>
	Займенник	<i>займ.;</i> <i>вказ.;</i> <i>пит.;</i> <i>особ.;</i> <i>означ.</i>
	Дієслово	<i>недок.;</i> <i>док.;</i> <i>акт.;</i> <i>безос.</i>
	Дієприкметник	<i>дієпр.</i>
	Прислівник	<i>присл.</i>
	Прийменник	<i>прійм.</i>
	Сполучник	<i>спол.</i>
	Частка	<i>част.</i>
	Вигук	<i>виг.</i>

№	Структурний елемент	Формальна ознака структурного елемента
		Ознака частини мови у правій частині
	Іменник	Абстр. ім.; Дія за знач.; Дія і стан за знач.; Властивість за знач.
	Прикметник	Прикм. до; Який; Стос. до; Одержаний.
	Дієприкметник	Дієпр.
	Прийменник	у знач. <i>прійм.</i>
	Прислівник	Присл. до.
4	Права частина, початок	Крапка, пробіл, цифра «1», крапка (прямий шрифт); крапка, пробіл, літера (українська, прямий шрифт, верхній регістр); двокрапка; Відсильне слово « <i>див.</i> » (курсив).
5	Стильові та граматичні позначки	Початок: текст після цифри (курсив). Закінчення: текст (прямий шрифт).
6	Тлумачення, початок	Крапка, пробіл, літера (українська, прямий шрифт, верхній регістр).
7	Тлумачення, закінчення	Текст (курсив), крапка, пробіл; квадратна відкривальна дужка (прямий шрифт);
8	Літературні джерела	Відкривальна дужка, текст, закривальна дужка (прямий шрифт).
9	Перелік атрибутів джерел	Через кому. Попередній атрибут закінчується крапкою або українською літерою (прямий шрифт).
10	Відтінок	Подвійна скісна риска «//» (прямий шрифт).
11	Фразеологізм	Один зі спеціальних символів (◇, △, ◆, ▲); абзац, одна літера (українська, напівжирний шрифт, верхній регістр), пробіл, літера (українська, нижній регістр); абзац, одна літера (українська, напівжирний шрифт, верхній регістр), літера (українська, нижній регістр).

№	Структурний елемент	Формальна ознака структурного елемента
12	Відношення	Абзац, одна літера (українська, світлий шрифт, верхній регістр), пробіл, літера (українська, нижній регістр), абзац включає слово «відношення».
13	Об'єднані тлумачення	Початок: двокрапка (прямий шрифт), літера «а», дужка (прямий шрифт). Наступна частина: крапка з комою (прямий шрифт), літера (українська) та дужка (прямий шрифт).
14	Усталені словосполучення	Двокрапка, один зі спеціальних символів (◇, △, ◆, ▲), одна літера (українська, напівжирний шрифт, верхній регістр), літера (українська, нижній регістр); двокрапка, один зі спеціальних символів (◇, △, ◆, ▲), одна літера (українська, напівжирний шрифт, верхній регістр), пробіл, літера (українська, нижній регістр).

У результаті конверсії весь текст СУМа з HTML-файлів автоматично перенесено до ЛБД, тобто здійснено парсинг Словника. Детальніше структуру ЛБД «СУМ» розглянуто у наступному підрозділі.

6.2. Структура і функції лексикографічної бази даних та інструментального комплексу «СУМ»

Структура ЛБД «СУМ» є мінімальною – вона дозволяє правильно відображати всі структурні елементи Словника, але водночас має можливість для розширення. Наприклад, ліва частина словникової статті у цій ЛБД не є структурованою, а записується єдиним блоком.

Структура ЛБД «СУМ» разом із клієнтською програмою дозволяє не тільки представити, але й візуалізувати представлення будь-якої словникової статті СУМа у вигляді дерева. При цьому значно спрощується доступ до її структурних елементів, а зв'язки між ними унаочнюються.

Чимало допоміжних елементів статті (термінальні символи, номери тлумачень, певні розділові знаки, спеціальні символи та шрифтові виділення) не потребують збереження у ЛБД, а можуть додаватися динамічно під час формування статті програмою виводу. Такі автоматичні операції допомагають уникнути багатьох помилок при редагуванні статей, а ймовірність помилкового введення елементів, які порушують структуру Словника, взагалі виключається. Процес редагування стає більш простим,

контрольованим та уніфікованим, легко здійснюються операції додавання, вилучення та виправлення елементів словникових статей.

Гнучкість структури ЛБД «СУМ» дозволяє у перспективі реалізувати те, що було принципово неможливим під час представлення його у формі послідовного тексту. Приміром, якщо фрагмент повторюється у різних статтях, то його можна зберігати в базі тільки один раз, але при цьому створити покажчики, які пов'язували б цей фрагмент з усіма відповідними статтями. Такий механізм разом з уже реалізованими функціями динамічного формування приводить до економії дискового простору.

У структурі ЛБД «СУМ» виділено сукупність таблиць та зв'язків між ними, зміст яких розкривається на рис. 6.2.

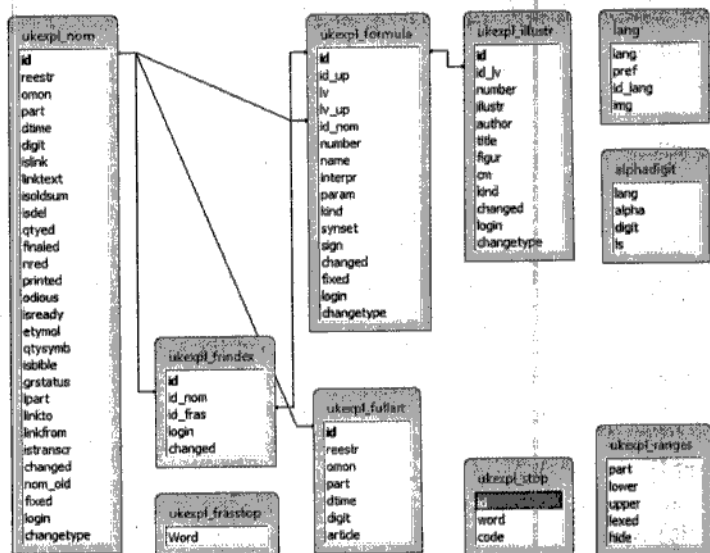


Рис. 6.2. Схема зв'язків між таблицями ЛБД «СУМ»

Далі наведено функціональне призначення та зміст полів окремих таблиць.

Таблиця мов і типу ЛБД (lang)

Таблиця передбачає можливість зберігання в одній фізичній базі даних ряду ЛБД (як тлумачного, так і інших типів) для різних світових мов:

- lang [nvarchar (255)] – довільний опис мови або ЛБД;
- pref [nvarchar (255)] – префікс для позначення таблиць у ЛБД;
- id_lang [int] – ідентифікатор мови.

Таблиця абетки та цифрових кодів (alphadigit)

Цифровий код використовується для сортування реєстру Словника. При цьому символи, які не мають такого коду, на послідовність сортування не впливають:

- lang [int] – код мови;
- alpha [nvarchar (10)] – літера або інший символ відповідної мови;
- digit [nvarchar (10)] – цифровий код символу;
- ls [tinyint] – порядок сортування літер із однаковим кодом (наприклад, великих і малих).

Таблиця діапазонів редагування (ukexpl_ranges)

Таблиця призначена для збереження організаційних даних із редагування ЛБД «СУМ» (організаційна структура проекту «СУМ-20» описана у підрозділі 6.5):

- part [nvarchar (5)] – номер Т-тому або частини Т-тому;
- lower [nvarchar (255)] – слово, яке є нижньою межею частини;
- upper [nvarchar (255)] – слово, яке є верхньою межею частини;
- lexed [nvarchar (255)] – прізвище та ініціали лексикографа або наукового редактора Т-тому або частини;
- hide [bit] – ознака, чи використовується даний діапазон.

Таблиця реєстрових слів СУМа (ukexpl_nom)

У цій таблиці зберігаються реєстрові слова СУМа, ліві частини словникових статей і ряд ознак, які стосуються словникових статей у цілому:

- id [int] – унікальний ідентифікатор статті;
- reestr [nvarchar (255)] – реєстрове слово СУМа;
- omon [int] – номер слова серед омонімів (0, якщо омонімії немає);
- part [int] – код частини мови реєстрового слова;
- dtime [datetime] – дата і час останньої правки статті;
- digit [nvarchar (255)] – цифровий код реєстрового слова, отриманий з використанням таблиці alphadigit;
- islink [int] – ознака, чи є стаття відсилною (1; при виводі позначається «div.»);
- linktext [nvarchar (255)] – текст відсилання;
- isoldsum [bit] – ознака, чи стаття отримана з 11-томного СУМа (1), чи є новою (0);
- isdel [bit] – ознака, чи статтю вилучено з ЛБД (1 або 0);
- qtyed [int] – кількість правок статті;
- finalized [bit] – ознака, чи була стаття погоджена з головним науковим редактором (1 або 0);

- **nred** [bit] – ознака, чи була стаття погоджена з науковим редактором (1 або 0);
- **printed** [bit] – ознака, чи була стаття відправлена на роздрукування (1 або 0);
- **odious** [bit] – ознака, чи належить стаття до проблемних (таких, які мають недоліки і підлягають подальшому перегляду й редагуванню, 1 або 0);
- **isready** [bit] – ознака, чи стаття пройшла видавниче редагування (1 або 0);
- **etymol** [nvarchar (4000)] – етимологія слова (зарезервовано, у СУМі не використовується);
- **qtysymb** [int] – кількість символів у статті;
- **grstatus** [nvarchar (255)] – ознака, чи звірено реєстрове слово з граматичним словником (1 або 0);
- **lpart** [nvarchar (4000)] – ліва частина словникової статті;
- **linkto** [nvarchar (255)] – зарезервовано;
- **linkfrom** [nvarchar (255)] – зарезервовано;
- **istranscr** [bit] – ознака, чи відредаговано транскрипцію (транскрипція є елементом інструментальної комп'ютерної системи, яка у паперовому варіанті не відтворюється);
- **changed** [tinyint] – ознака, чи статтю було змінено;
- **nom_old** [int] – ідентифікатор реєстрового слова у граматичному словнику;
- **fixed** [bit] – ознака, чи запис зафіксовано, тобто його не можна змінювати;
- **login** [int] – ідентифікатор користувача, який уносив останні зміни до статті;
- **changetype** [tinyint] – тип змін: якщо 0, то це системні зміни, історію яких не потрібно зберігати;

Таблиця проіндексована за полями: **id** (Unique), **reestr**, **omon**, **part**, **digit**.

Коди класів слів, що співвідносяться з певними частинами мови, наведено у табл. 6.2.

Таблиця тлумачних структурних елементів (**ukexpl_formula**)

Тут зберігаються елементи правих частин словникових статей, їх параметри та ознаки:

- **id** [int] – унікальний ідентифікатор тлумачного елемента;
- **id_up** [int] – ідентифікатор запису вищого рівня, з яким пов'язаний тлумачний елемент;
- **lv** [int] – рівень тлумачного елемента;

- **lv_up** [bit] – ознака того, чи тлумачний елемент пов'язаний безпосередньо із реєстровою одиницею або іншим тлумачним елементом;
- **id_nom** [int] – ідентифікатор статті, до якої належить тлумачний елемент;
- **number** [int] – порядковий номер тлумачного елемента в межах одного рівня;
- **name** [nvarchar (4000)] – назва тлумачного елемента;
- **interpr** [nvarchar (4000)] – текст формули тлумачення;
- **param** [nvarchar (1000)] – граматичні та стилістичні ремарки;
- **kind** [int] – тип тлумачного елемента;
- **synset** [int] – ідентифікатор синонімічного ряду, з яким пов'язаний тлумачний елемент;
- **sign** [int] – код типової формули тлумачення;
- **changed** [tinyint] – ознака, чи тлумачний елемент було змінено;
- **fixed** [bit] – ознака, чи запис зафіксовано, тобто його не можна змінювати;
- **login** [int] – ідентифікатор користувача, який уносив останні зміни до тлумачного елемента;
- **changetype** [tinyint] – тип змін: якщо 0, то це системні зміни, історію яких не потрібно зберігати.

Таблиця проіндексована за полями: **id** (Unique), **id_up**, **lv**, **lv_up**, **id_nom**, **number**.

Таблиця 6.2. Коды класів слів

1	Іменник	32	Вигук
3	Прикметник	33	Сполучник
4	Займенник	34	Частка
7	Дієприкметник	35	Прийменник
9	Числівник	41	Дієприслівник
13	Дієслово	21	Множинний іменник
31	Прислівник	22	Умовно-множинний іменник

Таблиця 6.3. Рівні структурних елементів словникових статей

0	Реєстрове слово	4	Значення словосполучення
1	Блок тлумачень	5	Відтінок тлумачення
2	Тлумачення слова	6	Дериват СС
3	СС		

Таблиця 6.4. Типи стійких словосполучень (СС)

Тип СС	Позначаються	Тип СС	Позначаються
Термінологічні	△	Фразеологізми	◇
Еквіваленти слова	○	Інші СС	Без позначки

Таблиця ілюстрацій (ukexpl_illustr)

У цій таблиці зберігаються приклади вживання реєстрових слів і стійких словосполучень з літературних та інших джерел:

- **id** [int] – унікальний ідентифікатор ілюстрації;
- **id_lv** [int] – ідентифікатор тлумачного елемента, з яким пов'язана ілюстрація;
- **number** [int] – номер ілюстрації в межах тлумачного елемента;
- **illustr** [nvarchar (4000)] – текст ілюстрації;
- **author** [nvarchar (255)] – автор джерела ілюстрації;
- **title** [nvarchar (255)] – назва джерела ілюстрації (зарезервовано для російського тлумачного словника);
- **figur** [bit] – ознака вживання реєстрового слова в ілюстрації в образному значенні;
- **cm** [bit] – ознака вживання реєстрового слова в ілюстрації у порівнянні;
- **kind** [int] – вид ілюстрації (зарезервовано для російського тлумачного словника);
- **changed** [tinyint] – ознака, чи ілюстрацію було змінено;
- **login** [int] – ідентифікатор користувача, який уносив останні зміни до ілюстрації;
- **changetype** [tinyint] – тип змін: якщо 0, то це системні зміни, історію яких не потрібно зберігати.

Таблиця проіндексована за полями: **id** (Unique), **id_lv**, **number**.

Таблиця фразеологічного індексу (ukexpl_findex)

Таблиця зберігає дані для побудови індексу СС СУМа:

- **id** [int] – унікальний ідентифікатор індексного запису;
- **id_nom** [int] – ідентифікатор статті, з якою пов'язаний запис;
- **id_fras** [int] – ідентифікатор фразеологізму;
- **login** [int] – ідентифікатор користувача, який уносив останні зміни до індексного запису;
- **changed** [tinyint] – ознака, чи запис було змінено;

Таблиця проіндексована за полями: **id** (Unique), **id_nom**, **id_fras**.

Інші допоміжні таблиці:

– **ukexpl_fullart** – таблиця для збереження словникових статей у вигляді неструктурованого HTML-тексту;

– **ukexpl_stop** – таблиця стоп-слів, які не потрібно індексувати під час суцільної повнотекстової індексації словникових статей СУМа;

– **ukexpl_frasstop** – таблиця слів, які не включаються до індексу СС.

Таблиці історії змін:

– **ukexpl_nom_log** – історія редагування статей;

– **ukexpl_formula_log** – історія редагування тлумачних елементів;

– **ukexpl_illustr_log** – історія редагування ілюстрацій;

– **ukexpl_frindex_log** – історія редагування фразеологічного індексу.

Ці таблиці мають такі додаткові поля:

– **id_log [int]** – унікальний ідентифікатор запису;

– **action [tinyint]** – тип зміни: 1 – update, 2 – insert, 3 – delete;

– **time_log [datetime]** – час, коли запис було змінено.

Інші поля мають такі самі імена й типи, як і відповідні поля таблиць **ukexpl_nom** та ін.

Запис до таблиць історії відбувається за відповідними тригерами. Для таблиць **ukexpl_nom** це тригери **ukexpl_nom_update**, **ukexpl_nom_insert**, **ukexpl_nom_delete**, для інших – за аналогією. Під час відповідної зміни даних тригер перевіряє значення поля **changetype**, і якщо воно не дорівнює 0 (тобто зміна виникла у ході редагування елемента статті, а не під час системних операцій), то попереднє значення запису зберігається в таблиці історії, у полі **action** установлюється тип зміни, а в полі **time_log** – поточний час.

Схема зв'язків таблиць ЛБД «СУМ» представлена на рис. 6.2.

Головною таблицею ЛБД є таблиця **ukexpl_nom**. Кожний її запис відповідає окремій словниковій статті СУМа і має у своєму складі атрибути, які характеризують статтю в цілому (реєстрове слово, омонім, ліва частина статті, відсилання, частина мови, кількість символів у статті, ознаки для контролю процесу редагування та ін.).

Таблиця **ukexpl_formula** має поля **id**, **id_up**, **lv_up**, **id_nom** та **lv**. Вони використовуються для встановлення зв'язків між записами таблиці різних рівнів; їх призначення наведено в описі таблиці. Наприклад, якщо в певному записі поле **lv** має значення 2, а **lv_up** – 0, то цей запис є тлумаченням, пов'язаним із реєстровим словом, **id** якого дорівнює значенню **id_up** для цього тлумачення. Якщо **lv_up** має значення 1, то запис пов'язаний з іншим тлумачним елементом (для тлумачення це має бути блок тлумачень) – тобто тлумачення є частиною цього елемента, на що вказує його значення **id_up**, яке дорівнює **id** відповідного запису блоку тлумачень.

При цьому **id_nom** як для тлумачення, так і для блоку тлумачень буде дорівнювати **id_reestr** слова статті, до якої вони належать.

Можливі зв'язки між рівнями тлумачних елементів (від верхнього рівня до нижчих) такі:

- 0: 1 (блоки тлумачень), 2 (тлумачення), 3 (СС);
- 1: 2 (тлумачення);
- 2: 4 (частини тлумачення), 5 (відтінки);
- 3: 4 (значення словосполучення), 6 (деривати);
- 4: 5 (відтінки);
- 5: 4 (частини відтінків).

Елементи рівнів 2, 4, 5, 6 також можуть мати ілюстрації, що реалізовано через зв'язок із таблицею **ukexpl_illustr**, а саме через поле **id_lv** у цій таблиці, що набуває значення **id** елемента, з яким пов'язана ілюстрація.

Таблиця **ukexpl_frindex** пов'язує таблиці реєстру та тлумачних елементів у такий спосіб, що кожному значенню **id_fras** (яке дорівнює певному **id** СС, тобто тлумачного елемента рівня 3) відповідає ряд значень **id_nom** – ідентифікаторів **id** відсильних слів, пов'язаних із даним СС.

Зв'язки між таблицями ЛБД «СУМ» відповідно до структури СУМа інтерпретуються таким чином:

- 1) стаття може складатися з одного або кількох блоків, тлумачень або СС. Вона може бути й відсильною і не мати у своєму складі жодного з перерахованих елементів;
- 2) блок може складатися з кількох тлумачень;
- 3) тлумачення може мати відтінки та частини тлумачення;
- 4) СС може мати кілька значень;
- 5) значення СС може мати відтінки;
- 6) тлумачення, відтінок, частина тлумачення, СС, значення СС можуть мати ілюстрації;
- 7) СС може мати відсилання на статті для слів, що входять до його складу.

Призначення окремих полів таблиці **ukexpl_formula** залежно від рівня тлумачного елемента мають певні відмінності:

- 1) блок не має формули тлумачення, але може мати ремарки, які записуються в поле **param**;
- 2) тлумачення може характеризуватися відношенням, код якого записується у поле **kind**. Для групи тлумачень з однаковим типом відношення, які мають послідовні номери, при формуванні статті тип відношення виводиться тільки для першого тлумачення з групи;

3) **CC** (тобто текст **CC**, до якого обов'язково входить реєстрове слово) завжди має заголовок, який записується у поле **name**. Також поле **kind** визначає тип **CC**, який позначається числом (див. табл. 6.4). Кожен тип характеризується відповідною поліграфічною позначкою при виведенні статті. **CC** однакового типу, що належать до спільного вищого рівня та мають послідовні номери, поєднуються у блок, позначка до якого виводиться тільки один раз.

Таблиці **ukexpl_formula** та **ukexpl_illustr** мають поле **number** для внутрішньої послідовної нумерації записів у межах рівня.

Визначено кілька правил зберігання окремих елементів словникових статей у таблицях ЛБД «СУМ», пов'язаних з особливостями формування статті для виведення на зовнішній пристрій:

— ліва частина зберігається у вигляді «**як є**» з усіма HTML-тегами, але без реєстрового слова на її початку. Реєстрове слово будується з полів **reestr** та **omon** таблиці **ukexpl_nom**, до нього автоматично додаються потрібні теги, а отриманий результат включається до лівої частини;

— якщо формула тлумачення закінчується крапкою, ця крапка не зберігається в базі даних, а додається автоматично програмою формування статті;

— ремарки формули тлумачення зберігаються без початкового та завершального тегів курсиву **<I>...</I>**, які при виведенні додаються автоматично;

— ілюстрації зберігаються без початкового та завершального тегів курсиву **<I>...</I>**; якщо вони завершуються крапкою або крапкою з комою, цей знак також не зберігається у базі даних;

— заголовки **CC** зберігаються без початкового та завершального тегів напівжирного шрифту **...**; якщо вони завершуються крапкою або крапкою з комою, цей знак також не зберігається у ЛБД.

Додавання нових записів до таблиць відбувається в такий спосіб. Значення поля **id** для нового запису формується автоматично засобами системи управління базами даних як автоінкремент, тобто на один більше поточного максимального значення. Інші умови розглянемо для окремих таблиць:

— таблиця **ukexpl_nom**. Значення **digit** визначається відповідно до реєстрового слова (див. опис табл.). Поле **part** заповнюється згідно з табл. 6.2. У поле **dtime** фіксується поточний час. Для більшості числових полів за замовчуванням визначено початкове значення, що дорівнює 0.

— таблиця **ukexpl_formula**. Поле **id_up** має дорівнювати **id** запису вищого рівня, з яким пов'язаний запис, що додається. Поле **lv** заповнюється так, як описано вище. Щоб визначити значення поля внутрішньої нумерації **number**, вибираються всі поля з цієї таблиці, які мають ті самі значення **id_up** та **lv_up**, тобто пов'язані з тим самим записом вищого рівня. З них обирається

найбільше значення **number** та до нього додається одиниця. Якщо запис потрібно не додати в кінець, а вставити між двома іншими, то здійснюється перенумеравання усіх записів, які повинні йти після того, що додається.

При видаленні цілої статті запис із таблиці **ukexpl_nom** насправді не видаляється з ЛБД, а значення ознаки **isdel** для нього встановлюється таким, що дорівнює одиниці.

При видаленні будь-якого фрагмента статті мають бути вилучені також усі пов'язані з ним елементи нижчих рівнів. Наприклад, при видаленні тлумачення відбувається також пошук і видалення усіх пов'язаних із ним відтінків та ілюстрацій.

6.3. Віртуальна лексикографічна лабораторія «СУМ-20»

Інструментальний комплекс «СУМ» спроектовано за системотехнікою віртуальної лексикографічної системи. Це поняття, уперше введене у монографії [276], визначено в такий спосіб. У процесі функціонування агрегованих лексикографічних систем у реальних соціотехнічних середовищах може відбуватися і їхня віртуалізація. У тому випадку, коли предметна галузь X – носій суперсистеми лексикографічних систем – набуває розподілених системних характеристик і параметризується певною структурованою множиною системних мережевих адрес (наприклад таких, що прийняті в Інтернеті), кожен елемент $x \in X$ постає функцією від певного кортежу адрес:

$$x = f(a_1 a_2 \dots a_n).$$

Тобто лексикографічна система $LSx[L]$ перетворюється на віртуальний об'єкт, розподілений у фізичному просторі, що зображається точкою (a_1, a_2, \dots, a_n) у просторі мережевих адрес. При цьому узгодження відповідних моделей даних на концептуальному та внутрішньому рівнях навіть не вимагається (хоча для зручності таке узгодження було б, звичайно, корисним). Необхідне лише узгодження на рівні зовнішніх моделей, зокрема на рівні мережевих протоколів, що забезпечує мінімальну цілісність віртуальної лексикографічної системи та можливість її ідентифікації як єдиного об'єкта. Така віртуальна лексикографічна система може бути використана як віртуальна лексикографічна лабораторія (ВЛЛ), що надає засоби для виконання спільних лексикографічних проєктів фахівцями різних територіально розподілених інституцій і навіть різних країн.

Однак об'єднання віртуальних лексикографічних систем у корпоративну мережу пов'язане зі значними матеріальними витратами, а саме створенням комунікаційної інфраструктури з високою пропускну здат-

ністю, вартість якої у наш час досить велика. Адже корпоративна мережа – це складна система, яка містить тисячі різноманітних компонентів: комп'ютери різних типів, починаючи з персональних та закінчуючи дата-центрами і великими кластерами, системне та прикладне програмне забезпечення, мережеві адаптери, концентратори, комутатори та маршрутизатори, фізичну мережу транспорту даних. Останнім часом простежується зміна технології автоматизованої обробки корпоративної інформації, пов'язана зі зростанням популярності Інтернету, який усе більше використовується як засіб забезпечення професійної взаємодії фахівців, що працюють над розв'язанням спільних проблем.

Таким чином, існуюча комунікаційна інфраструктура Інтернету цілком спроможна забезпечити необхідну пропускну здатність і нормальну роботу ВЛЛ у межах її запланованих функцій. Це може досягатися за допомогою різних програмно-технологічних інструментів, зокрема шляхом застосування системотехніки так званої сервіс-орієнтованої архітектури (Service-Oriented Architecture – SOA).

Основні аспекти створення розподілених систем типу ВЛЛ такі:

- гетерогенність сучасних інформаційних систем;
- обмін метаданими між системами різних виробників;
- обмін даними між системами, які суттєво різняться використанням різних форматів представлення даних;
- великі об'єми даних, які циркулюють між системами;
- гарантування доставки повідомлень;
- маршрутизація повідомлень й адресація «кінцевих точок»;
- координація процесів;
- безпека сервісної взаємодії.

Принципи, на яких будується ВЛЛ «СУМ», полягають у тому, що вона розробляється на базі Web-сервісів (коротко WS) – одній із реалізацій SOA-застосувань, що позиціонуються як універсальна технологія зв'язування суттєво різнорідних систем. В основі технології лежить декілька стандартів: XML для опису даних, SOAP для передачі інформації з одних систем на інші, WSDL для опису сервісів (у тому числі задання типів вхідних і вихідних даних) і UDDI для зберігання й надання WSDL-описів за запитом.

Цих стандартів достатньо для створення відносно нескладних систем. Однак мінімально нетривіальні рішення (котрі зазвичай і потрібні в корпоративному середовищі) вимагають забезпечення таких аспектів, як гарантована асинхронна доставка повідомлень, керування транзакціями, шифрування даних, які циркулюють між системами, забезпечення їхньої недоторканності. Усі ці аспекти так чи інакше споріднені з WS, і зараз ак-

тивно розробляються інструменти надбудови різних специфікацій, що дає можливість вписати їх у світ WS.

Визначення та огляд засобів і технологій проектування мережевих застосувань подано у працях [310; 228].

ВЛЛ «СУМ», побудовану на основі викладених принципів, створено у середовищі розробки Microsoft Visual C# 2005 Professional Edition. Вона працює під операційною системою Microsoft Windows XP/2003, Vista або 7 під управлінням Microsoft.NET Framework версії 3.5.

Комплекс має багаторівневу архітектуру: сервер бази даних відповідає за зв'язок із ЛБД і за функції отримання та збереження даних; сесійний сервер установлює сеанси роботи для окремих користувачів, керує привілеями та встановлює рівні доступу; клієнтська програма надає інтерфейс користувача, що створює можливість виконувати лексикографічну і редакційну роботу, а також ряд інших функцій.

Програмний комплекс орієнтовано на роботу в мережі (як локальний, так і глобальний, оскільки використання технології створення розподілених сервіс-орієнтованих систем Windows Communication Foundation (WCF) для взаємодії між окремими рівнями комплексу дає можливість його ефективного функціонування у середовищі Інтернету), де багато користувачів одночасно мають доступ до ЛБД «СУМ». При цьому залежно від привілеїв користувачі можуть отримувати доступ до всієї бази або її частини, можливість редагування статей або тільки їх перегляду тощо.

Програмний комплекс ВЛЛ «СУМ» у локальному та віддаленому режимах забезпечує виконання таких функцій:

- автентифікація та авторизація користувачів;
 - внесення нової словникової статті до ЛБД «СУМ»;
 - видалення словникової статті з ЛБД «СУМ»;
 - редагування словникової статті у ЛБД «СУМ»;
 - забезпечення модифікації структури словникової статті певної лексикографічної системи ВЛЛ і створення нових лексикографічних систем;
 - динамічне формування словникових статей ЛБД «СУМ» у друкарському форматі;
 - виконання різного роду вибірок і створення підсистем;
 - аналіз даних ЛБД «СУМ»;
 - додання нового користувача до ВЛЛ «СУМ»;
 - видалення користувача з ВЛЛ «СУМ»;
 - керування правами доступу користувачів ВЛЛ «СУМ».
- Вхід користувача ВЛЛ «СУМ» до системи здійснюється через підсистему доступу, вікна якої зображено на рис. 6.3.

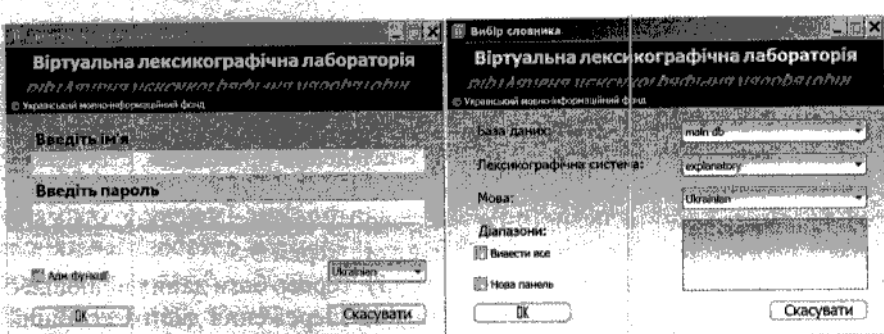


Рис. 6.3. Вікна входу до ВЛЛ

Функціональні можливості програмного комплексу ВЛЛ «СУМ» показано на рис. 6.4.

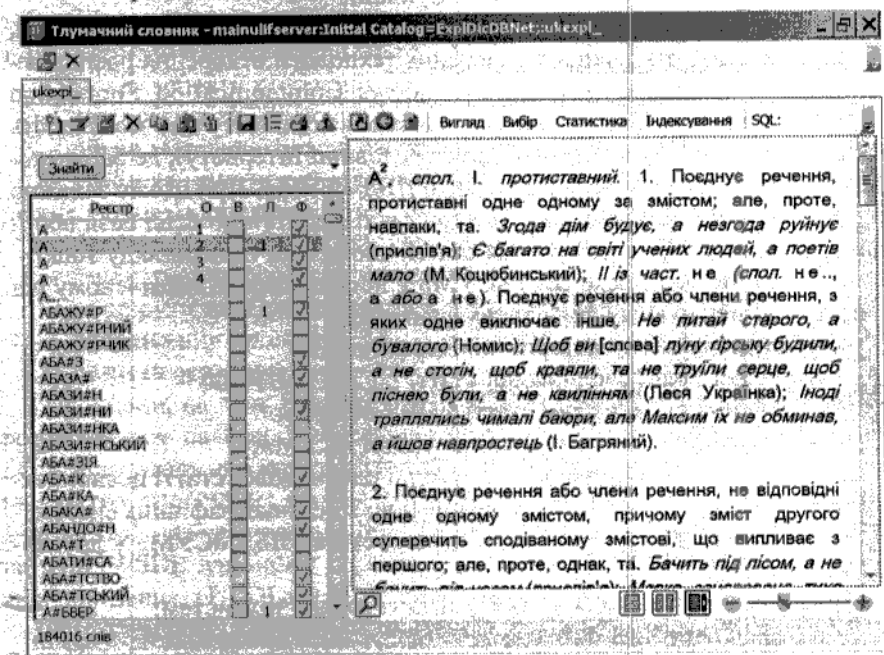
















Рис. 6.4. Головне вікно редагування лексикографічної інформації ВЛЛ «СУМ»

У головному вікні редагування лексикографічної інформації ВЛЛ «СУМ» ліворуч по вертикалі відображається реєстр словника з такими стовпчиками:

- Реєстр – заголовне слово словникової статті;
- O – номер омоніма;
- B – ознака, чи опрацьована стаття видавничим редактором;
- L – ознака, чи була стаття додана або змінена;
- T – ознака, чи стаття опрацьована технологом.

Праворуч відображається текст словникової статті.

Кнопки в панелі інструментів мають таке призначення:

-  «Додати» – ввести нову словникову статтю;
 -  «Редагувати» – редагувати словникову статтю;
 -  «Редагувати копію» – створити або редагувати альтернативну редакцію словникової статті;
 -  «Вилучити» – вилучити словникову статтю;
 -  «Скопіювати» – створити копію вибраної статті;
 -  «Історія редагування» – переглянути історію змін у словниковій статті;
 -  «Інші редакції статті» – відобразити вікно вибору альтернативних редакцій;
 -  «Записати статті» – зберегти вибрані статті або діапазон статей у HTML-форматі;
 -  «Записати фразеологічний індекс» – зберегти у HTML-форматі файл фразеологічного індексу для вибраного діапазону;
 -  «Роздрукувати статтю»;
 -  «ЛЯП!» – позначити вибрану статтю як проблемну (або зняти позначку, якщо в неї внесено правки);
 -  «Перейти до слова» – звернення до статті через слово, яке вибране у тексті поточної словникової статті;
 -  «Повернутися» – повернення до попередньої статті;
 -  «Форум» – перейти до веб-форуму укладачів і технологів СУМа;
 - «SQL:» – фільтрація реєстру Словника за довільною умовою у форматі SQL, яка вводиться в сусідньому рядку.
- Також через меню програми доступні такі функції:
- установлення параметрів перегляду словникової статті (шрифт, позначки наголосу, необхідність виведення ілюстрації, виділення кольором нових та змінених фрагментів статті і т. ін.);
 - вибір діапазонів реєстру за виконавцями – лексикографами, науковими редакторами і т. ін.;
 - фільтрація реєстру за різними критеріями (статті, наявні або відсутні в 11-томному СУМі, додані та змінені статті, за частиною мови, за відсутністю тлумачень та ілюстрацій, за стилістичними ремарками, типовими формулами тлумачення і т. ін.);

– отримання статистичних показників (кількість усіх, доданих та змінених статей, значень або ілюстрацій у будь-якому діапазоні реєстру, кількість знаків у статтях, діаграми за кількістю стилістичних ремарок і типових формул тлумачення).

– індексування словникових статей у потрібному діапазоні.

Вікно редагування словникової статті має вигляд, показаний на рис. 6.5.

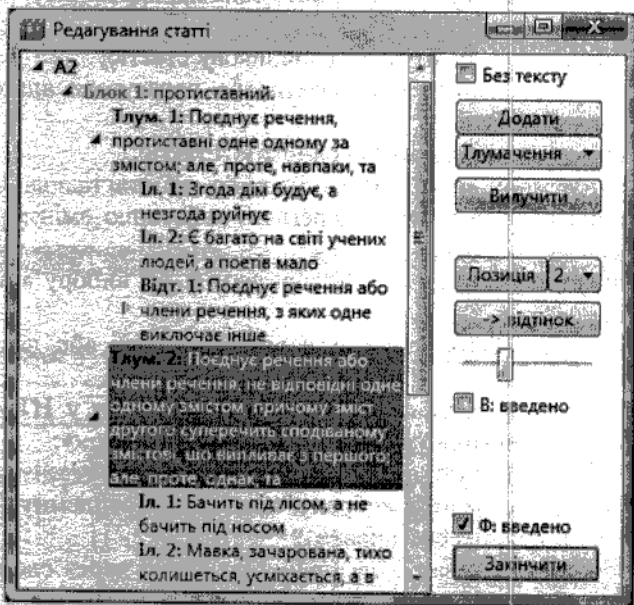


Рис. 6.5. Вікно редагування словникової статті

Структура словникової статті має вигляд дерева, в якому представлені:

- «Блок» – блок тлумачень;
- «Тлум.» – тлумачення;
- «Відт.» – відтінок тлумачення;
- «Слсп.» – стійке словосполучення;
- «Знач.» – значення словосполучення;
- «*)гл.» – частина тлумачення або відтінку;
- «Дер.» – дери́ват фразеологізму;
- «Іл.» – текстова ілюстрація.

Після вибору структурного елемента доступні такі функції:

«Додати» – додати структурний елемент. Можна додати елемент того ж рівня (наприклад, на рис. 6.5 додано тлумачення після «Тлум. 2», при цьому

номери наступних тлумачень збільшаться на 1). Також можна додати елемент нижчого рівня (наприклад, до вибраного «Тлум. 2» – відтінок, ілюстрацію і т. ін., які будуть додані до вже існуючих – «Відт. 1» або «Лл. 3» відповідно);
«Вилучити» – вилучити вибраний структурний елемент та всі, пов'язані з ним;

«Позиція» – переставити елемент статті на позицію того ж рівня, яка вибирається зі списку (наприклад, «Тлум. 1» на рис. 6.5 можна переставити на 2-ге або 3-тє місце);

«-> відтінок» або «-> тлумачення» – перетворити тлумачення на відтінок або навпаки;

«В: введено» – ознака, чи опрацьована стаття видавничим редактором;

«Т: введено» – ознака, чи опрацьована стаття технологом;

«Закінчити» – завершити редагування статті.

Також можна встановити режим відображення дерева статті у спрощеному вигляді («Без тексту») та обрати бажаний розмір шрифту у вікні.

Вікно редагування слова (рис. 6.6) дає можливість увести реєстрове слово, номер його омоніма та частину мови.

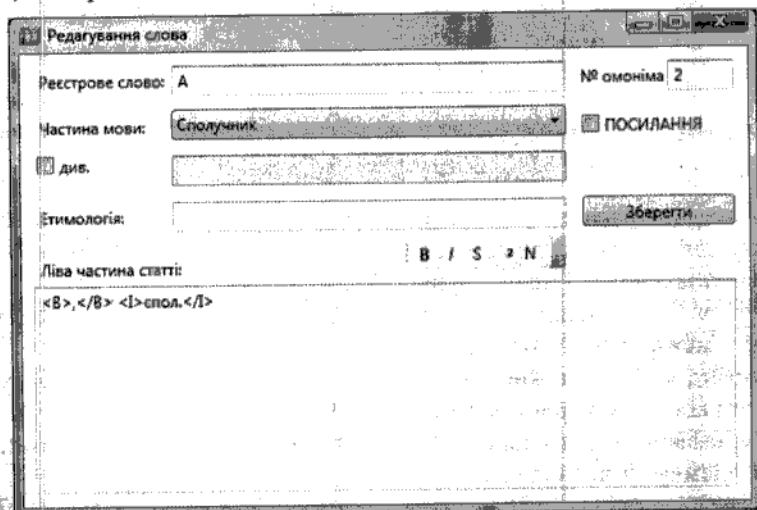


Рис. 6.6. Вікно редагування слова

Якщо стаття відсильна, то встановлюється відповідна ознака і вводиться реєстрове слово, на яке потрібне відсилання. Тут є і поле для редагування лівої частини словникової статті та поле етимології слова (зарезервовано для наступних редакцій СУМа).

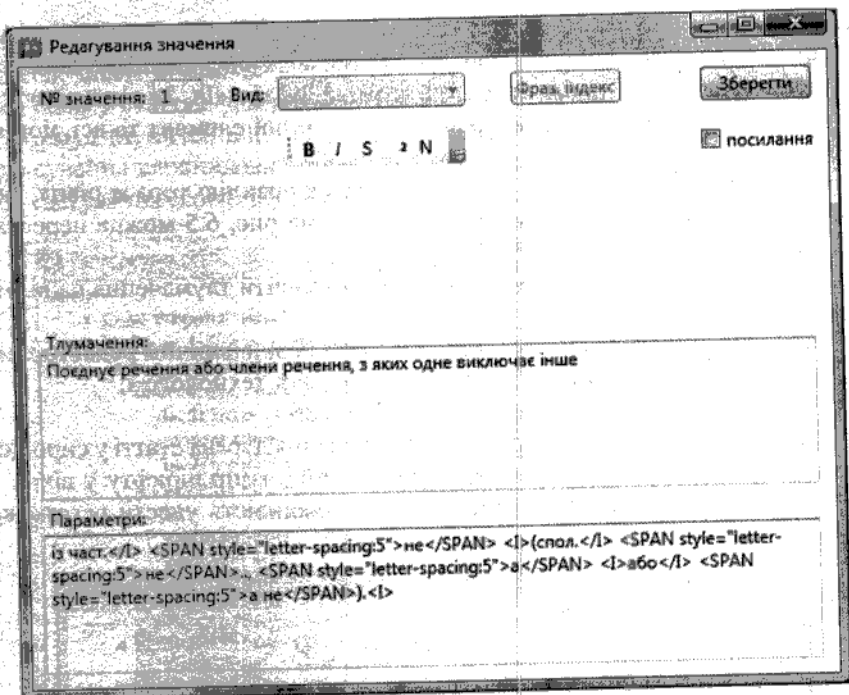


Рис. 6.7. Вікно редагування словникової дефініції

Вікно редагування словникової дефініції (рис. 6.7) набуває подібного вигляду для всіх елементів статей, які мають формулу тлумачення. У ньому представлено такі поля і функції:

- «№» – порядковий номер тлумачення (або іншого елемента статті) у межах одного рівня;
- «Вид.» – залежно від елемента статті може характеризувати його тип або не використовуватися;
- «Назва.» – для словосполучень та їх дериватів подає текст, для інших елементів не використовується;
- «Тлумачення.» – текст тлумачення;
- «Параметри.» – граматичні позначки, стилістичні ремарки;
- «Зберегти» – зберегти зміни в базі даних;
- B / S * N – панель редагування шрифтових тегів у тексті. Теги встановлюються для виділеного фрагмента тексту.

Перед збереженням у ЛБД виконується перевірка коректності введеного тексту на правильність розставлення HTML-тегів. Якщо знайдено

помилки, то збереження не дозволяється, доки вони не будуть виправлені. Перевірку можна відключити натисненням відповідної комбінації клавіш.

Для редагування ілюстрацій у програмі є вікно, яке має подібний вигляд та набір функцій.

Серед інших допоміжних функцій програми редагування зазначимо:

- редагування фразеологічного індексу (додавання або вилучення реєстрових слів, пов'язаних із вибраним фразеологізмом);
- редагування стоп-слів, які не підлягають індексації або повинні бути проіндексовані автоматично;
- автоматизоване введення похідних слів для доданого слова залежно від його частини мови;
- блокування окремих діапазонів реєстру від редагування та зняття цього блокування.

Застосування інструментального комплексу «СУМ» у формі ВЛЛ продемонструвало високу ефективність у процесі укладання СУМ-20, дозволивши сформувати корпус Словника за відносно короткий час і надавши можливість безперервного його розвитку та вдосконалення шляхом моніторингу мовної динаміки української мови.

6.4. Лінгвістичний корпус

Необхідним елементом комп'ютерної технології нового СУМа став лінгвістичний корпус, який прийшов на зміну традиційній лексичній картотеці і відіграв роль одного з основних джерел текстово-ілюстративного матеріалу до СУМ-20. Оскільки корпусна лінгвістика є досить новим напрямом лінгвістичного дослідження й досі у вітчизняній лексикографічній практиці не застосовувалася, розглянемо основні її аспекти і будову та функції Українського національного лінгвістичного корпусу (УНЛК), створеного в Українському мовно-інформаційному фонді, а також його використання при укладанні СУМ-20.

У загальному розумінні лінгвістичний корпус — це представлений в електронному вигляді, великий за обсягом, уніфікований, структурований, розмічений і філологічно компетентний масив мовних даних, доповнений системою керування — універсальними програмними засобами для пошуку різноманітної лінгвістичної інформації [283:11–17]. Обсяги мовного матеріалу, який залучається до мовознавчого дослідження, комплексність, оперативність опрацювання зазначеного матеріалу та можливість прямого доступу до великої кількості лінгвістичних фактів — це ті переваги, які надає лінгвістичний корпус досліднику.

Сервіс УНЛК акумулює функціональні можливості сервісів інших інформаційно-лексикографічних систем, а саме — електронної бібліотеки, граматичного словника, інтегрованої лексикографічної системи «Словники України», а також функцій повнотекстового пошуку. До всіх перерахованих систем розроблено окремі клієнтські програми, що забезпечують інтерфейси роботи з ними. УНЛК надає єдиний інтерфейс до всіх цих сервісів, створюючи ефект роботи з єдиною системою.

Кожен зазначений сервіс відповідає за функціональність певної лексикографічної системи окремо, а взаємодія цих сервісів разом у програмному комплексі УНЛК створює необхідне середовище обробки та збереження інформаційних ресурсів.

У системі УНЛК виділяються дві основні підсистеми: бібліографічна та повнотекстова. Бібліографічна частина являє собою електронну бібліотеку — колекцію електронних ресурсів, яка є основою для розробки будь-якого корпусу. Тому бібліографічна підсистема служить інструментом для збирання, збереження, моделювання й використання природномовної інформації в цифровому вигляді і становить невід'ємну частину УНЛК. Вона виконує роль багатофункціональної інформаційної системи, яка акумулює інформацію різних видів. Ефективність електронної бібліотеки можлива лише за умови використання чіткої та прозорої схеми представлення метаданих об'єктів збереження. Питання стандартизації опису даних розглядається як у межах кожної установи, так і на державному та міжнародному рівнях. Це питання є складним — адже сучасні розвинені стандарти бібліографічного опису налічують тисячі полів. Їх семантика не завжди однозначна, що стає причиною колізій при інтеграції ресурсів. Проте ми не виключаємо необхідності дотримання відповідності встановленим стандартам. Однак, як переконує практика, при створенні колекцій об'єктів різномірної природи для різних масивів об'єктів використовуються різні підмножини відповідного стандарту — формат опису даних.

Розглянемо це питання з позиції системотехніки організації електронної бібліотеки. Якщо внутрішню структуру збереження метаданих тісно пов'язують зі встановленим форматом опису даних, виникає проблема масштабованості системи. Будь-яка зміна формату (навіть незначна) викликає лавиноподібну зміну всього програмного комплексу, що, у свою чергу, призводить до необхідності розробки окремих програмних систем для різних колекцій або спричиняє неперервний процес розробки та супроводження програмної системи. Розв'язанням цієї проблеми може стати відмова від прив'язування структури бази даних збереження метаданих об'єктів електронної бібліотеки до конкретних даних. Оперування понят-

тими архітектури системи переносить відповідність даних до сутностей із рівня логіки роботи з даними на рівень логіки застосувань. Подібний підхід дозволив спроектувати базу даних у такий спосіб, щоб вона служила універсальним сховищем збереження метаданих різної природи, а її структура не залежала від формату опису об'єктів збереження.

Формат опису об'єктів, що зберігаються в базі даних, визначає адміністратор системи. Для цього створено спеціальну підсистему (її вхідне вікно показано на рис. 6.8), в якій адміністратор може визначити всі можливі характеристики, за якими будуть описуватися об'єкти. Вони використовуються при внесенні метаданих про об'єкт.

Редагування таблиці властивостей						
Зберегти		Додати		Видалити		
Назва характеристики	Код характеристики	Тип даних	Словник	Група характеристик	Можливість повторення	
Назва видання	200-A Осередок	Символьний	Назва видання	Назва видання	<input checked="" type="checkbox"/>	
Тип назви видання	200-B Загальна	Числовий	Тип назви	Назва видання	<input checked="" type="checkbox"/>	
Відомість, що	200-B	Символьний			<input checked="" type="checkbox"/>	
Тип автора	200-C	Числовий	Тип автора	Автор	<input checked="" type="checkbox"/>	
Прізвище та ініціали		Числовий	Автори	Автор	<input type="checkbox"/>	
Відомість про	710-A Назва	Символьний			<input checked="" type="checkbox"/>	
Нова видання	104-A	Числовий	Мова	Нова видання	<input checked="" type="checkbox"/>	
Тип мови видання	104-A Мова тексту	Числовий	Нова видання	Нова видання	<input checked="" type="checkbox"/>	
Шлях до файлу	332	Шлях до файлу			<input checked="" type="checkbox"/>	
Тип файлу	332	Числовий	Тип файлу		<input checked="" type="checkbox"/>	
Ід-е файлу	332	Символьний			<input checked="" type="checkbox"/>	
Видання/час	202-B	Числовий	Видання/час		<input checked="" type="checkbox"/>	
Рік видання (3)	210-B	Числовий		Рік видання	<input checked="" type="checkbox"/>	
Рік видання (10)	210-D	Числовий		Рік видання	<input checked="" type="checkbox"/>	
Рік видання (10)	210-B	Символьний		Рік видання	<input checked="" type="checkbox"/>	
Місце видання	210-A	Числовий	Місце видання		<input checked="" type="checkbox"/>	
Відомість про	202-A	Символьний			<input checked="" type="checkbox"/>	
Серія	225-A Назва серії	Числовий	Серія		<input checked="" type="checkbox"/>	
Тема	102-A	Числовий	Тема		<input checked="" type="checkbox"/>	
Жанр	102-A-B	Числовий	Жанр		<input checked="" type="checkbox"/>	
Ключові слова	010-A Тематичні	Числовий	Ключові слова		<input checked="" type="checkbox"/>	
Номер тому (3)	207-A	Числовий		Том	<input checked="" type="checkbox"/>	
Номер тому (10)	207-B	Числовий		Том	<input checked="" type="checkbox"/>	

Рис. 6.8. Підсистема редагування формату опису об'єктів збереження

У режимі редагування адміністратор може визначити назву характеристики, що виводиться на екран, зазначити інформацію про додаткові домовленості, відповідність стандарту, додати коментарі, які допоможуть користувачеві точніше з'ясувати зміст тієї чи іншої назви. Адміністратор визначає також типи значень, яких може набувати конкретна характеристика. Відповідно до викладеного автоматично генеруються форми для введення значень, будуються пошукові запити, внесена інформація перевіряється на коректність введення та при обробці даних викликаються відповідні функції.

У системі для метаданих визначено п'ять можливих типів: символічний, числовий, дата, шлях до файлу та блоб-об'єкт. Як видно з назв типів, метаданими можуть виступати не лише текстові та числові дані, але й посилання на документи, що зберігаються у файльовій системі сервера або ж будь-які двійкові дані, занесені в базу (наприклад, фото, аудіо, відео).

Певні характеристики можуть набувати значення з обмеженої множини значень. Для полегшення роботи з такими характеристиками в системі передбачена можливість підключення певних словників значень. У адміністратора системи є можливість створити новий або ж відредагувати існуючий словник. Він може бути двох видів: закритий або відкритий. Словники відкритого типу можуть бути відредаговані бібліографом системи, а закриті словники доступні для редагування лише адміністратору. Приміром, до характеристики «мова видання» під'єднано словник, де визначено назви мов за відповідним стандартом.

Наступний параметр характеристики, який адміністратор може змінювати, – це група. Деякі підмножини характеристик можуть бути пов'язані в логічні групи – наявність однієї з характеристик в описі об'єкта вимагає обов'язкової наявності всіх інших характеристик з цієї групи. Форми для заповнення значень елементів групи генеруються окремо для кожної з груп. Останній із параметрів характеристики – можливість повторення, що визначає унікальність наявності характеристики в межах опису одного об'єкта.

Наступним етапом налаштування системи є створення профілів характеристик, під якими розуміється певна підмножина характеристик із загальної множини. Кожний користувач має можливість вибору (визначення) певного профілю (профілів) для пошуку та заповнення даних.

До функцій адміністрування системи входить редагування інформації про користувачів і визначення рівня доступу до системи. Таких рівнів у системі передбачено чотири: читач, бібліограф, адміністратор і адміністратор системи.

Функції бібліографа системи полягають у наповненні бази об'єктів збереження. Наприклад, для УНЛК – це внесення літературного джерела до бази, заповнення необхідної метайнформації джерела, процес підготовки тексту до повнотекстової індексації. Якщо в системі працює одночасно кілька бібліографів, є можливість переглянути статистику поповнення бази даних одним із бібліографів за певний період.

Пошук за бібліографічними реквізитами (див. рис. 6.9) є функцією, доступною широкому колу користувачів. Пошукове поле користувач обирає самостійно з полів, що увійшли до пошукового профілю. Якщо це

текстове поле, то користувач вводить інформацію, якщо ж поле має обмежену множину значень, то користувач обирає пошукове значення з відповідного словника. Для розширеного пошуку використовується об'єднання результатів за логікою «і» чи «або». Додатковими параметрами пошуку є ознака наявності у джерела його повного тексту для повнотекстового індексу, а також ознаки індексації.

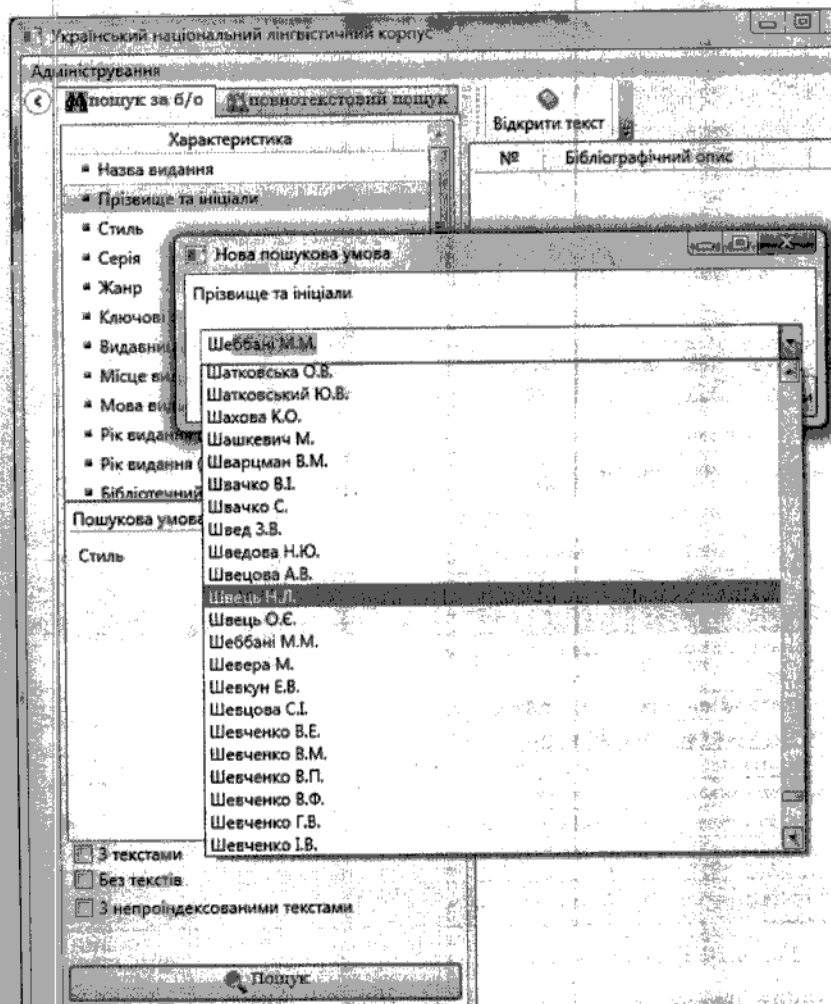


Рис. 6.9. Пошук за елементами бібліографічного опису

Результати пошуку представляються у вигляді списку бібліографічних описів (див. рис. 6.10). Користувач має можливість переглянути повний список бібліографічних реквізитів для кожного об'єкта, переглянути ресурс (повний текст за його наявності), а також записати в окремий файл результати пошуку.

Відкрити повний текст Запис БО до файлу Кошик об'єктів збереження

Нечуй-Левицький І. Без пуття : Оповідання. - <http://www.utoronto.ca/elul/Nechui/BezPuttia/index.html> - К. : Наук. думка, 1966.

Нечуй-Левицький І. Гетьман Іван Виговський : Роман / Елібіотека з української літератури для школярів – CD-ROM 4x - К. : Студія "Негоціант", 2000. - (Золота скарбниця України). - ISBN 966-7423-50-6

Нечуй-Левицький І. Дві москвки : Оповідання. - http://lib.ru/SU/UKRAINA/NECHUIJ_LEVIC_KIJ/zmoskovk.txt

Нечуй-Левицький І. Кайдашева сім'я. Хмари : Повість - К. : Наук. думка, 2001. - 501 с. - (Шк. б-ка). - ISBN 966-00-0676-4

Нечуй-Левицький І. Київські прохачі : Оповідання. -

Знайдено: 511 Час завантаження: 1 с

Завантажити прикріплений файл Розмір контексту: 200

Бібліографічні дані Статистичні дані Контексти

Назва характеристики	Значення характеристики	At
Назва видання	Т. 1: Прозові твори.	
Тип назви видання	Основна назва книги	
Тип автора	автор	
Прізвище та ініціали	Нечуй-Левицький І.	
Відомості про відповід...	http://www.utoronto.ca/elul/Nechui/	
Мова видання	Ukrainian	
Тип мови видання	Мова видання	
Шлях до файлу	book\dl000944.rar	
Тип файлу	файл в текстовому форматі	

Запис створено: Терещенко Лариса Олександрівна

Рис. 6.10. Результати пошуку за елементами бібліографічного опису

Друга можливість доступу до об'єктів – це повнотекстовий пошук (див. рис. 6.11).

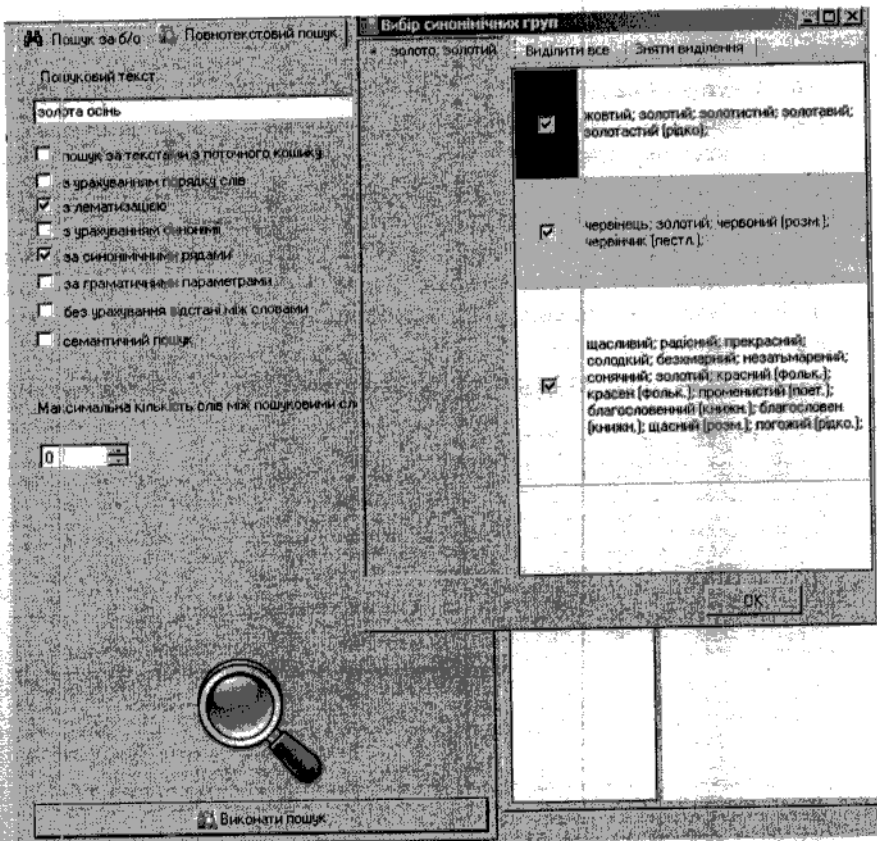


Рис. 6.11. Вікно повнотекстового пошуку

Користувач вводить пошукову фразу, задає максимально бажану кількість слів між пошуковими та обирає додаткові параметри повнотекстового пошуку, а саме:

- урахування порядку слів;
- пошук у певній підмножині об'єктів;
- використання процедури лематизації (зведення текстового слова до початкової форми);
- використання синонімічної лексикографічної бази даних;
- використання певних синонімічних рядів;
- вибір граматичних параметрів для кожного слова, що входить у пошуковий фрагмент.

Результатом повнотекстового пошуку є список бібліографічних описів. Але на відміну від пошуку за бібліографією користувач отримує прямий доступ до кожної локалізації пошукового фрагмента в тексті, тобто до всіх контекстів, які містять пошуковий фрагмент. Обравши джерело, користувач може переглядати контексти, де пошуковий фрагмент виділено червоним кольором. Розмір (довжину) контексту можна змінювати. Кожне слово з контексту (не лише пошукове) з'єднане зі словниками: граматичним, тлумачним і словником синонімів. Для подальшої обробки всі контексти або контексти певного джерела можуть бути записані в HTML-файл із зазначенням джерела контексту, часу створення та пошукової фрази (див. рис. 6.12).

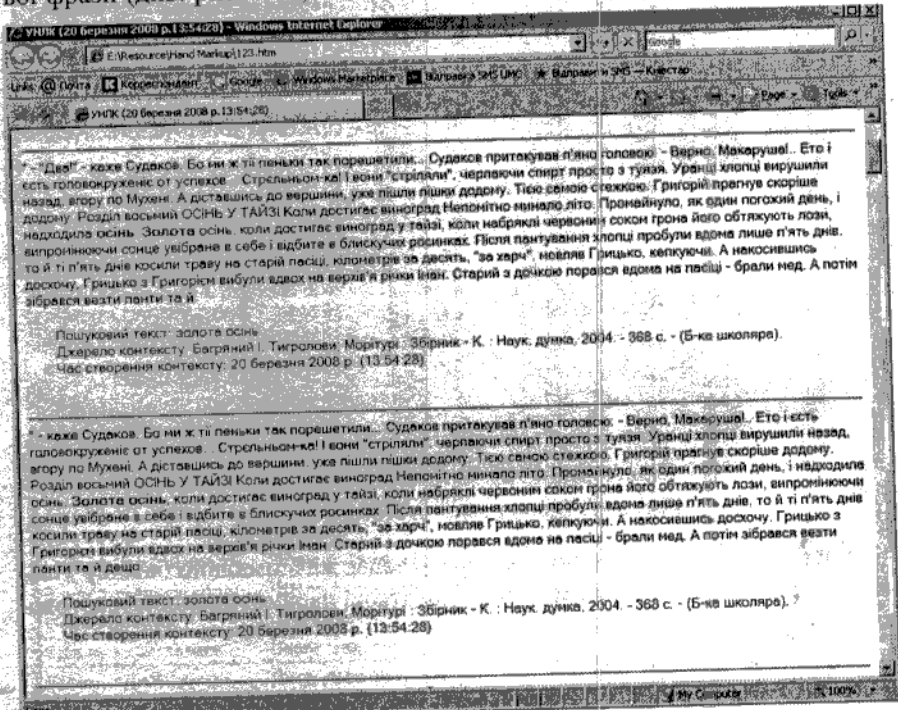


Рис. 6.12. Результати повнотекстового пошуку

Ще одним засобом для створення користувацьких наборів даних у системі є так званий кошик, тобто тимчасове сховище. Користувач може відібрати об'єкти збереження, які його цікавлять, з різних пошукових запитів і зберегти образ даних для подальшої роботи в інших сеансах.

Користуючись таким інструментом, дослідник може відібрати джерела певного автора, стилю чи жанру і працювати лише з цією частиною корпусу; фактично, ця процедура дає можливість виділити із загального корпусу свій власний підкорпус, орієнтований на розв'язання особистих завдань. Такий підхід дозволив відмовитися від ідеї створення еталонного корпусу, оскільки ідея еталонності, з нашого погляду, є занадто суб'єктивною і до того ж відсекає можливість дослідження засобами корпусу нееталонних, аномальних мовних явищ. Не зовсім вдалою, на нашу думку, є ідея виваженого, або збалансованого корпусу, що передбачає реалізацію попередньо заданих співвідношень об'ємів літературних джерел, які належать до різних стилів, жанрів, епох і т. ін. Адже універсальних критеріїв виваженості або збалансованості немає, і всі спроби у цьому напрямі відбивають лише особисті погляди їх авторів.

Словник | Статистичні дані

Тип статистики: Кількість джерел за авторами

Сортування за: алфавітом частотністю

Кількість елементів: 2953

Нечуй-Левицький І.	21
Даррелл Д.	13
Савицький С. П.	13
Широков В. А.	13
Гончар О. Т.	12
Франко І.	11
Вовчок М.	9
Загребельний П.	9
Кащенко А.	9
Мирний П.	9
Грищенко О. М.	8
Ваграний І.	7
Бельдія А. Я.	7
Левкай В. С.	7
Русанівський В. М.	7
Величко С. М.	6
Винниченко В.	6
Довженко О.	6
Карпенко-Карий І.	6
Мушкетер Ю.	6
Бережний В.	5
Білецький В. С.	5

Словник | Статистичні дані

Тип статистики: Реєстри з лематизацією

Сортування за: алфавітом частотністю

Кількість елементів: 149301

агломеративний	8
агломерація	38
аглофабрика	2
аглоупіат	11
аглоупіативний	10
аглоупіація	23
аглоупіати	1
аглоупіоген.	1
аглоупіуватися	1
агнеса	28
агнесць	166
агні-Йога	2
агнозія	3
агностика	4
агностицизм	8
агностицистичний	4
агон	134
агоніка	1
агонізувати	14
агоніст	113
агонія	130
агор	2

Рис. 6.13. Приклад статистичних даних

Таким чином, реалізовані в УНЛК функції кошика дозволяють породжувати в режимі реального часу та в рамках єдиного лінгвістичного корпусу віртуальні підкорпуси, які моделюють ті або інші лінгвістичні ефекти, із забезпеченням кожному дослідникові можливостей реалізації його власних уявлень про сталонність, виваженість і збалансованість.

Для об'єктів, занесених до кошика, дозволено стандартні функції доступу до тексту, перегляду бібліографічного опису, а також додаткові функції перегляду статистичних параметрів і запису вихідних текстів на клієнтську машину.

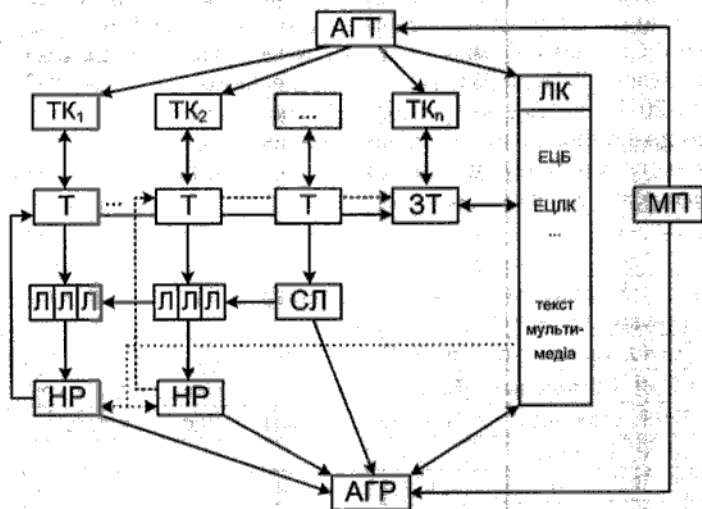
Для користувача доступні статистичні дані корпусу (див. рис. 6.13). Виділено декілька видів статистики як за бібліографічними даними, так і за даними повнотекстового пошуку. Передбачено два види сортування для статистичних даних: за алфавітом або за частотністю.

На сьогодні УНЛК містить 4259 літературних джерел обсягом 1,09 Гб, що налічують понад 64 млн слововживань.

6.5. Організація лексикографічного процесу при укладанні СУМ-20

Лексикографічну роботу зі створення СУМ-20 організовано за принципом лексикографічного конвєса [187; 150; 188; 282; 301–303; 285; 286; 305]. Згідно з цим принципом вона зосереджується в описаному вище технологічному середовищі «СУМ», де весь лексикографічний процес розподіляється за послідовностями взаємопов'язаних і взаємозалежних кроків із одночасним виконанням певних операцій над лексикографічними об'єктами. Основним лексикографічно-технологічним кластером – науково та організаційно визначеним квантом лексикографічної інформації – тут виступає так званий технологічний том (Т-гом). Він являє собою певний (загалом незв'язний) фрагмент Словника, закріплений за певним колективом лексикографів, який при укладанні першої редакції складався з п'яти осіб: науковий редактор технологічного тому – керівник лексикографічної групи, три лексикографи і комп'ютерний технолог. У такий спосіб сформовано дев'ять Т-томів і відповідно – дев'ять лексикографічних груп, за кожною з яких закріплено певні діапазони опрацювання Словника (їх наведено у табл. 6.5). Складено поточні перспективні плани опрацювання лексикографічної інформації (розширення реєстру, дослідження семантики, формування словникових дефініцій, добір текстово-ілюстративного матеріалу тощо), хід виконання яких контролювався менеджментом проекту «СУМ». Кожен фрагмент Словника, опрацьований лексикографічною гру-

пою, подавався головному науковому редакторові, а після урахування його зауважень – до науково-видавничого редагування. Схему, яка діяла при укладанні першої редакції СУМ-20, у другій редакції було дещо модифіковано. На рис. 6.14 наведено фрагмент загальної організаційно-технологічної схеми створення лексикографічної системи (ЛС) СУМ у такому середовищі.



АГР – адміністрація головного редактора
АГТ – адміністрація головного технолога
ЕЦБ – електронно-цифрова бібліотека
ЕЦЛК – електронно-цифрова лексична картотека
ЗТ – зведення комп'ютерної інформації

Л – лексикограф
ЛК – лінгвістичний корпус української мови
МП – менеджер проекту
НР – науковий редактор
СЛ – спеціальна лексикографія
ТКі – технологічні комплекси ЛС СУМ
Т – комп'ютерний технолог

Рис. 6.14

Наведений фрагмент ілюструє організацію взаємодії менеджменту проекту ЛС СУМ (В. В. Манако та Ю. В. Леонтєв), адміністрації головного наукового редактора (головний науковий редактор – В. М. Русанівський, його заступники, наукові редактори і лексикографи) та адміністрації головного технолога (керівник О. Г. Рабудець, головний програміст К. М. Якименко, комп'ютерні технологи С. І. Волошина, І. М. Кондрачук, М. П. Слабковська).

Основними кроками лексикографічної роботи зі створення СУМ-20 стали такі:

- парсинг СУМ-11 та формування ЛБД «СУМ»;
- укладання на основі цієї ЛБД протягом 2001–2006 років першої редакції СУМ-20 (далі СУМ-20-1) у вигляді дев'яти технологічних томів ЛС-СУМ загальним обсягом близько 174 тис. словникових статей;
- створення на основі дев'яти видавничо-технологічних томів першої редакції нового Словника вісімнадцяти видавничо-технологічних томів другої його редакції (далі СУМ-20-2) та формування першого тому СУМ-20-2 (А–Б).

У СУМ-20-1, як показано у табл. 6.5, специфіковано 63 частини, які, відповідно до затвердженої організаційно-технологічної схеми, обробляли 34 лексикографи, 9 наукових редакторів, головний науковий редактор і 4 технологи технологічних томів.

Таблиця 6.5

№ Т-тому	Науковий редактор Т-тому	Специфікація Т-тому (ТТ)
1	О. О. Тараненко	А, Д, Е, Є, Ж, И, І, І, Й, М, Р – розвиватися
2	С. Я. Єрмоленко	Б, Л, П – переступник, розвиднитися – ряхтливий
3	Н. Г. Озерова	В, перестьобувати – підслизовий, У
4	В. Б. Фридрак	Г, Г, К, О
5	Л. Л. Шевченко	З, підсліджувати – пов'ючити
6	В. В. Чумак	Н, пов'язаний – покалічтися, Т
7	В. М. Брицин	по-калмицьки – промовницький
8	Г. М. Ярун	промовчаний – пропурхнути, С
9	М. М. Пешак, В. О. Винник	пропуск – п'ять, Ф, Х, Ц, Ч, Ш, Щ, Ъ, Ю, Я

Лексикографи технологічних томів: Г. Н. Горюшина, Н. Є. Лозова, І. А. Самоїлова (перший Т-том); С. П. Бирик, С. Я. Єрмоленко, Н. М. Неровня, Г. М. Сюта (другий Т-том); І. С. Гнатюк, Г. І. Лиса, Л. М. Стоян (третій Т-том); Л. В. Мовчун, О. М. Тищенко, Т. О. Федоренко, В. Б. Фридрак (четвертий Т-том); М. П. Годована, Т. Я. Марченко, О. А. Стишов, Л. Л. Шевченко (п'ятий Т-том); В. М. Білоноженко, І. О. Борисюк, І. А. Казімірова, В. В. Чумак (шостий Т-том); В. М. Брицин, І. Ю. Гальчук, В. В. Дятчук, О. І. Скопненко (сьомий Т-том); А. К. Мойсієнко, Л. О. Ставицька, Г. М. Ярун, Н. Ю. Ясакова (восьмий Т-том); С. О. Вербич, Л. Е. Пшенична, Н. М. Сухарина (Заїка) (дев'ятий Т-том).

На базі дев'яти технологічних томів сформовано дев'ять видавничо-технологічних томів, упорядкованих за абеткою. Протягом 2004–2005 років виконано науково-видавниче редагування першого видавничо-технологіч-

ного тому СУМ-20-1 (А – горішок; перша редакція). Науковий редактор – Н. Г. Озерова, лексикографи – Г. Н. Горюшина (А, 422 стор.), Г. М. Стюга (Б, 806 стор.), І. С. Гнатюк (В – від'єднувати, 1315 стор.), Г. І. Лиса (від'ємний – в'яльце, 586 стор.), І. В. Кришан (віяний – в'ятір, 504 стор.), Л. В. Мовчун (Г – говорильня, 437 стор.), В. Б. Фридрак (говорити – горішок, 126 стор.). Відповідальний видавничий редактор – А. Я. Бельдій, керівник групи видавничих редакторів – Л. В. Туник, видавничі редактори першої редакції першого видавничо-технологічного тому – Г. С. Балабан, О. А. Дітель, Л. М. Кудрявкіна, Є. І. Мазніченко, Н. М. Максименко, О. В. Осадча, Л. С. Шорубалка, М. А. Щипська, Л. К. Ярошевська.

На початку 2007 року на базі дев'яти видавничо-технологічних томів першої редакції СУМ-20 специфіковано 18 видавничо-технологічних томів і поставлено завдання – створити другу редакцію СУМ-20. Друга редакція лексикографічно опрацьовується переважно співробітниками Українського мовно-інформаційного фонду Національної академії наук України. Відзначимо, що вихідний текст першого тому другої редакції сформовано лексикографами Фонду в режимі прямого доступу до лексикографічної бази даних (а не в режимі опосередкованого доступу через комп'ютерних технологів), що дозволило різко підвищити продуктивність та якість їхньої роботи. Саме тоді розпочалася промислова експлуатація ВЛЛ «СУМ».

В укладанні першого тому СУМ-20-2 (А–Б; друга редакція) взяла участь Л. Л. Шевченко (А – аннамі), В. В. Чумак (аннібалів – аятола, Б – балухи), Г. М. Ярун (балцанка – бенетити), І. В. Шевченко (бенефіс – божеволіти), О. В. Бугаков (боженька – бусольний), В. М. Білоноженко (бустер – бязь). Лексикографічну працю з укладання усталених словосполучень у СУМ-20-2 здійснювала В. М. Білоноженко. Обов'язки наукового редактора першого тому виконувала Н. Г. Озерова. Науковими консультантами СУМ-20-2 були: з граматичних питань О. О. Тараненко, з питань стилістики та культури мови – С. Я. Ермоленко, із видавничо-редакційної роботи – Л. В. Туник. Редакторську роботу виконувала Н. В. Щербакова. Технологічну підтримку лексикографічної роботи здійснювали О. Г. Рабулець (керівник технологічної групи до 2010 року), К. М. Якименко (головний технолог-програміст), Н. М. Сидорчук (лінгвістичний корпус) та комп'ютерні технологи С. І. Волошина, І. М. Кондрачук, М. П. Слабковська. Індекс першого тому СУМ-20-2 формували М. Ю. Кригін та І. В. Остапова. Менеджерами другої редакції були В. В. Манако та Ю. В. Леонтєв (планування і контроль робіт, розроблення технологічних схем, організація робіт з розроблення та проведення автоматизованого тестування ЛБД ЛС СУМ, взаємодія з видавництвом). Видавничі редактори першого тому СУМ-20-2 – працівники Державного науково-виробничого підприємства «Видавництво “Наукова думка” НАН України» – Н. М. Максименко, Г. С. Балабан, К. С. Чайка. Загальне наукове керівництво здійснював В. А. Широков.

Рецензенти першого тому: академік НАН України Г. П. Півторак; доктор філологічних наук, професор М. П. Кочерган; доктор філологічних наук, професор Л. І. Мацько; доктор філологічних наук, професор А. К. Мойсієнко; доктор філологічних наук, професор М. Я. Плющ.

У процесі оновлення фундаментальних лексиконів завжди виникають питання: наскільки нова лексикографічна праця є насправді новою? Чи достатньо мовних інновацій уведено до словника, щоб уважати результат виконаної роботи справді новим словником? Ще глибші та складніші проблеми виникають при бажанні встановити кількісні характеристики новизни та оригінальності лексикографічної праці. Слід зауважити, що досі чітких критеріїв новизни словників у світовій лексикографічній теорії не вироблено. За лексикографічною традицією (яка передається лише в усній формі – від одного покоління лексикографів до наступного) прийнято, що словник є новим, якщо його зміст відрізняється від прототипу не менше як на 15%. Але як вираховувати відсотки – це питання вже не уточнюється і, в свою чергу, породжує багато нових запитань. Зрозуміло, що автори СУМ-20 також неодноразово про це дискутували.

Застосування комп'ютерних технологій лексикографування при укладанні СУМ-20 дозволило вперше в українському мовознавстві розпочати роботу з фундаментальної лексикографічної статистики. Наведемо деякі статистичні параметри СУМ-11 та СУМ-20, аналіз яких дає можливість здійснити своєрідну кваліметрию словникових текстів, на основі чого читач може зробити аргументований висновок про те, що СУМ-20 є цілком новою лексикографічною працею, оскільки він радикально відрізняється від СУМ-11. Статистичні дані стосуватимуться діапазону першого тому (А–Б) СУМ-20, оскільки він є повністю завершений, у той час як до наступних діапазонів ще будуть внесені зміни, і цей процес триватиме аж до виходу в світ останнього тому.

Наприклад, у табл. 6.6 наведено статистичні дані про відносні обсяги діапазону першого тому (А–Б) для першої (СУМ-20-1) та другої (СУМ-20-2) редакцій:

Таблиця 6.6²

Назва показника	СУМ-20-1	СУМ-20-2	Збільшення у СУМ-20-2 порівняно з СУМ-20-1, %
Словникові статті	9665	11531	19,31
Знаки у словникових статтях	2841457	3921919	38,02
Ілюстрації	18250	26406	44,69

² Тут і далі всі зіставлення статистичних показників (СУМ-11, СУМ-20-1, СУМ-20-2) стосуються тільки діапазону А–Б.

З табл. 6.6 видно, що друга редакція першого тому СУМ-20 значно відрізняється від першої: реєстр збільшився майже на 20%, майже удвічі виросла кількість ілюстрацій, на 38% збільшився загальний обсяг лексикографічного матеріалу. Отже, і за традиційними лексикографічними поглядами другу редакцію СУМ-20 можна вважати новим словником навіть порівняно з його першою редакцією. При цьому кількісні показники, звичайно, зовсім не відбивають тих фактичних змін, які були зроблені лексикографами у процесі дослідження розвитку семантичної структури лексики та системи словосполучень.

У табл. 6.7 наведено докладні статистичні дані щодо відносних пропорцій лексикографічних параметрів діапазону першого тому СУМ-11 і СУМ-20 (друга редакція).

Таблиця 6.7

Назва показника	СУМ-11	СУМ-20-2	Збільшення у СУМ-20-2 порівняно із СУМ-11, %
Словникові статті	6303	11531	82,94
Знаки у словникових статтях	1786334	3921919	119,55
Тлумачення реєстрових слів, у т. ч.:	9577	14346	49,80
– значень;	8495	13079	53,96
– відтінків значень	1082	1267	17,10
Ілюстрації	11604	26406	127,56
Усталені словосполучення (з тлумаченнями та посланнями):	643	2252	250,23
– СС;	439	522	18,91
– еквіваленти слів;	0	6	–
– термінологічні словосполучення;	51	479	839,22
– фразеологізми	153	1245	713,73

З табл. 6.7 помітно, що обсяг діапазону А–Б у СУМ-20 порівняно із СУМ-11 збільшився більш ніж удвічі. Обсяг реєстру майже подвоївся. На 127% виросло число ілюстрацій, що є свідченням великої роботи з добору текстово-ілюстративного матеріалу, оскільки в 11-томнику значна кількість реєстрових слів та їх значень не проілюстрована. У новому ж Словнику такі випадки поодинокі й умотивовані: ілюстрації не подаються лише тоді, коли вони справді не потрібні. Особливо значне зростання спостерігається у системі словосполучень. По-перше, тут з'явився якісно новий елемент – еквіваленти слова, що є свідченням розвитку мовознавчої теорії за період, який пройшов від часу видання 11-томника. Велике – понад 700% – зростання числа фразеологізмів пояснюється значною роботою, яку виконала українська лексикографія упродовж останніх 30 років. Зокрема, до нового тлумачного Словника практично повністю інтегровано матеріали фунда-

ментальних словників: «Фразеологічного словника української мови» у двох книгах, виданого у 1993 році, та «Словника фразеологізмів української мови», що вийшов друком у серії «Словники України» у 2003 році. Особливо новий Словник поповнився термінологічними словосполученнями – їх загальна кількість у першому томі виросла більше як у вісім разів, що є свідченням мовної динаміки, спричиненої науково-технічними революціями, які відбулися і відбуваються починаючи з кінця 70-х років минулого століття, а також змінами у суспільно-політичній системі.

Саме цими факторами пояснюється динаміка вживання галузевих ремарок у СУМ-20 порівняно із СУМ-11, яку відображено в табл. 6.8 (наведено лише ті ремарки, збільшення яких перевищує 100 %). При цьому слід зауважити, що ряд ремарок був відсутній у діапазоні А–Б 11-томника (ветеринарія, друкарська справа, релігія та ін.).

Таблиця 6.8

Ремарка	Повна назва	СУМ-11	СУМ-20-2	Збільшення у СУМ-20-2 порівняно із СУМ-11, %
<i>ав.</i>	авіація	7	15	114,29
<i>анат.</i>	анатомія	8	31	287,50
<i>антр.</i>	антропологія	2	4	100,00
<i>архт.</i>	архітектура	11	26	136,36
<i>астр.</i>	астрономія	5	11	120,00
<i>бібл.</i>	біблійне слово (вислів)	0	29	–
<i>біол.</i>	біологія	28	217	675,00
<i>біохім.</i>	біохімія	6	21	250,00
<i>бот.</i>	ботаніка	11	34	209,09
<i>буд.</i>	будівельна справа	6	20	233,33
<i>вет.</i>	ветеринарія	0	83	–
<i>військ.</i>	військова справа	7	14	100,00
<i>геогр.</i>	географія	1	2	100,00
<i>геод.</i>	геодезія	2	7	250,00
<i>геол.</i>	геологія	13	47	261,54
<i>гірн.</i>	гірництво	3	6	100,00
<i>екол.</i>	екологія	0	29	–
<i>екон.</i>	економіка	9	43	377,78
<i>ел.</i>	електрика	1	6	500,00
<i>етн.</i>	етнографія	6	12	100,00
<i>жарг.</i>	жаргонне слово (вислів)	0	21	–
<i>зоол.</i>	зоологія	1	17	1600,00

Ремарка	Повна назва	СУМ-11	СУМ-20-2	Збільшення у СУМ-20-2 порівняно із СУМ-11, %
<i>інформ.</i>	інформатика	0	53	
<i>іст.</i>	історія; історизм	39	144	269,23
<i>кул.</i>	кулінарія	2	4	100,00
<i>лінгв.</i>	лінгвістика	38	134	252,63
<i>літ.</i>	літературознавство; література	20	52	160,00
<i>мат.</i>	математика	8	55	587,50
<i>мед.</i>	медицина	47	302	542,55
<i>мет.</i>	металургія	7	16	128,57
<i>метеор.</i>	метеорологія	3	8	166,67
<i>мист.</i>	мистецтво	2	8	300,00
<i>мор.</i>	морська справа	6	45	650,00
<i>муз.</i>	музика	14	47	235,71
<i>палеонт.</i>	палеонтологія	0	17	—
<i>поет.</i>	поетичне слово (вислів)	17	42	147,06
<i>полігр.</i>	поліграфія	0	17	—
<i>політ.</i>	політика	6	14	133,33
<i>псих.</i>	психологія	9	20	122,22
<i>радіо</i>	радіотехніка	1	2	100,00
<i>рел.</i>	релігія	0	37	—
<i>рел.- церк.</i>	релігійно-церковне слово (вислів)	0	32	—
<i>спец.</i>	спеціальне слово	64	247	285,94
<i>спорт.</i>	спортивний термін	10	20	100,00
<i>театр.</i>	театральний термін	1	2	100,00
<i>текст.</i>	термін текстильної промисловості	0	6	—
<i>фарм.</i>	фармацевтика	6	71	1083,33
<i>фіз.</i>	фізика	25	95	280,00
<i>фізіл.</i>	фізіологія	6	16	166,67
<i>філос.</i>	філософія	14	40	185,71
<i>фін.</i>	фінансова справа	10	79	690,00
<i>фольк.</i>	фольклор	3	6	100,00
<i>фот.</i>	фотографія	0	1	—
<i>хім.</i>	хімія	26	206	692,31
<i>юр.</i>	юридичний термін	9	31	244,44

У табл. 6.9 показано пропорції лексики за частинами мови у реєстрах СУМ-11 і СУМ-20-2 цього ж діапазону.

Таблиця 6.9

Частини мови	СУМ-11	СУМ-20-2	СУМ-11, %	СУМ-20-2, %
Іменник	3558	6981	56,45	60,54
Прикметник	1595	2997	25,31	25,99
Дієслово	510	665	8,09	5,77
Прислівник	389	559	6,17	4,85
Дієприкметник	64	109	1,02	0,95
Займенник	12	20	0,19	0,17
Числівник	0	1	0,00	0,01
Інші	175	199	2,78	1,73

З цієї таблиці простежуємо, що пропорції за окремими лексико-граматичними класами суттєво не помінялися. Звертає на себе увагу деяке збільшення (близько 4,5%) у СУМ-20-2 порівняно із СУМ-11 відносної кількості іменників на тлі приблизно такого ж (3,89%) зменшення відносної кількості дієслів. Це можна пояснити деяким зростанням (а саме на 3%) термінологічної лексики у СУМ-20 порівняно із СУМ-11, яка представлена переважно іменниками.

Проте порівняння цих самих пропорцій, вирахованих на всьому масиві СУМ-11, дає іншу картину: іменник – 41,44%; дієслово – 30,93%; прикметник – 16,52%; дієприкметник – 5,45%; прислівник – 4,99%; займенник – 0,08%; числівник – 0,06%; інші – 0,53%. Це пояснюється особливостями діапазону слів, що починаються на літери А і Б, зокрема відносною бідністю дієслів на фоні явного домінування іменника. У наступних таблицях (6.10 та 6.11) наведено аналогічну статистику за частинами мови в межах окремих літер А та Б.

Таблиця 6.10 (літера А)

Частини мови	СУМ-11	СУМ-20-2	СУМ-11, %	СУМ-20-2, %
Іменник	1172	2805	60,73	63,92
Прикметник	450	1065	23,32	24,27
Дієслово	103	192	5,34	4,38
Прислівник	71	124	3,68	2,83
Дієприкметник	40	76	2,07	1,73
Займенник	8	10	0,41	0,23
Числівник	0	0	0,00	0,00
Інші	86	116	4,46	2,64

Таблиця 6.11 (літера Б)

Частина мови	СУМ-11	СУМ-20-2	СУМ-11, %	СУМ-20-2, %
Іменник	2386	4176	54,57	58,46
Прикметник	1145	1932	26,19	27,05
Дієслово	407	473	9,31	6,62
Прислівник	318	435	7,27	6,09
Дієприкметник	24	33	0,55	0,46
Займенник	4	10	0,09	0,14
Числівник	0	1	0,00	0,01
Інші	88	83	2,01	1,16

У табл. 6.12 та 6.13 представлено відносні пропорції лексикографічних параметрів за літерами А та Б.

Таблиця 6.12 (літера А)

Назва показника	СУМ-11	СУМ-20-2	Збільшення у СУМ-20-2 порівняно з СУМ-11, %
Словникові статті	1931	4388	127,24
Знаки у словникових статтях	511791	1379297	169,50
Тлумачення реєстрових слів, у т. ч.:	2460	5212	111,87
– значень;	2189	4865	122,25
– відтінків значень	271	347	28,04
Ілюстрації	2679	8201	206,12
Усталені словосполучення, у т. ч.:	130	406	212,31
– СС;	79	90	13,92
– еквіваленти слів;	0	1	–
– термінологічні словосполучення;	28	243	767,86
– фразеологізми	23	72	213,04

Таблиця 6.13 (літера Б)

Назва показника	СУМ-11	СУМ-20-2	Збільшення у СУМ-20-2 порівняно із СУМ-11, %
Словникові статті	4372	7143	63,38
Знаки у словникових статтях	1277118	2542622	99,09
Тлумачення ресстрових слів, у т. ч.:	6035	9134	51,35
– значень;	5224	8214	57,24
– відтінків значень	811	920	13,44
Ілюстрації	8925	18205	103,98
Усталені словосполучення (з тлумаченнями та посиланнями):	513	1846	259,84
– СС;	360	432	20,00
– еквіваленти слів;	0	5	–
– термінологічні словосполучення;	23	236	926,09
– фразеологізми	130	1173	802,31

У табл. 6.14 наведено дані про обсяги ілюстрацій у СУМ-11 і СУМ-20-2.

Таблиця 6.14

Діапазон	Обсяг ілюстрацій у СУМ-11 у знаках та % до всього діапазону	Обсяг ілюстрацій у СУМ-20-2 у знаках та % до всього діапазону	Збільшення обсягу ілюстрацій СУМ-20-2 порівняно із СУМ-11, %
А-Б	899 855 (50,3%)	2 361 712 (60,2%)	162,46
Літера А	239 941 (46,9%)	841 111 (60,98%)	250,5
Літера Б	659 914 (51,7%)	1 520 601 (59,8%)	130,4

Лексикографічну статистику плануємо наводити в кожному томі – як для даного тому, так і для суми попередніх томів включно. У такий спосіб читач матиме змогу відстежувати статистичні параметри процесу лексикографічного розвитку Словника української мови.

Кожний новий том Словника після його видання плануємо виставити в українському сегменті Wiktionary з метою надання вільного доступу користувачам української мови для Wiki-лексикографування. Увесь словниковий масив СУМ-20 досяжний уже зараз у режимі спеціального доступу за адресою: <http://lcorp.ulif.org.ua/xclientinfo/>, де зазначені умови користування ним.

ЛІТЕРАТУРА

Словники

1. Англійсько-український словник з математики та кібернетики / Уклад. Є. В. Мейнарович, М. І. Кратко.— К. ; Ірпінь : ВТФ «Перун», 2010.— 568 с.
2. Англійсько-український словник з програмування і математики / Уклад. М. І. Кратко, М. І. Кікець, Є. В. Мейнарович та ін.— Луцьк : Надстир'я, 1998.— 640 с.
3. Англо-український математичний словник / Ред. І. А. Черненко.— К.: НВП «Дидактик», 1993.— 224 с.
4. Англо-український фразеологічний словник / Уклад. К. Т. Баранцев.— К. : Рад. шк.— 2005.— 1051 с.
5. Бацевич Ф. С. Словник термінів міжкультурної комунікації.— К. : Довіра, 2007.— 205 с.
6. Бибик С. П., Єрмоленко С. Я., Пустовіт Л. О. Словник епітетів української мови.— К. : Довіра, 1998.— 431 с.
7. Білорусько-український словник / За ред. Г. П. Півторака.— К.: Довіра, 2006.— 723 с.
8. Великий англо-український словник / Упоряд. М. Г. Зубков.— Донецьк: Сталкер; К. : ЗАТ «НКП», 2006.— 1152 с.
9. Великий зведений орфографічний словник сучасної української лексики / Уклад. і голов. ред. В. Т. Бусел.— К. ; Ірпінь : ВТФ «Перун», 2003.— 887 с.
10. Великий російсько-український політехнічний словник / За ред. О. С. Благовещенського.— К. : Вид. дім «Чумацький шлях», 2002.— 749 с.
11. Великий тлумачний словник. Сучасна українська мова від А до Я / Упоряд. А. П. Загнітко.— Донецьк : ТОВ ВКФ «БАО», 2008.— 704 с.
12. Великий тлумачний словник української мови / Упоряд. Т. В. Ковальова.— Харків : Фоліо, 2005.— 767 с.
13. Годована М. П. Словник-довідник назв осіб за видом діяльності.— К. : Наук. думка, 2009.— 176 с.

14. Головащук С. І. Орфографічний словник складних слів української мови.— К.: Наук. думка, 2008.— 512 с.
15. Головащук С. І. Російсько-український словник сталих словосполучень.— К.: Наук. думка, 2004.— 640 с.
16. Головащук С. І. Словник-довідник з правопису та слововживання / За ред. В. М. Русанівського.— К.: Наук. думка, 1989.— 830 с.
17. Головащук С. І. Словник-довідник з українського літературного слововживання.— К.: Наук. думка, 2004.— 448 с.
18. Головащук С. І. Словник наголосів.— К.: Наук. думка, 2002.— 319 с.
19. Гринчишин Д. Г., Капелюшний А. О., Сербенська О. А., Терлак З. М. Словник-довідник з культури української мови.— Львів: Фенікс, 1996.— 368 с.
20. Гринчишин Д. Г., Сербенська О. А. Словник паронімів української мови.— К.: Освіта, 2008.— 256 с.
21. Даль В. И. Толковый словарь живого великорусского языка. Современное написание: В 4 т.— М.: АСТ-Астрель, 2002.
22. Данилюк І. Г. Великий сучасний російсько-український українсько-російський словник фразеологізмів та сталих виразів.— Донецьк: ТОВ ВКФ «БАО», 2009.— 640 с.
23. Демська О. М., Кульчицький І. М. Словник омонімів української мови.— Львів: Фенікс, 1996.— 223 с.
24. Етимологічний словник української мови: В 7 т.— К.: Наук. думка, 1982–2006.— Т. 1–5.
25. Ефремова Т. Ф. Современный толковый словарь русского языка: В 3 т.— М.: АСТ-Астрель, 2006.
26. Ефремова Т. Ф. Толковый словарь словообразовательных единиц русского языка.— М.: АСТ-Астрель, 2005.— 636 с.
27. Ефремова Т. Ф. Толковый словарь служебных частей речи русского языка.— М.: АСТ-Астрель, 2004.— 814 с.
28. Єрмоленко С. Я., Єрмоленко В. І., Ленець К. В., Пустовіт Л. О. Новий російсько-український словник-довідник.— К.: Довіра, 1996.— 797 с.
29. Жайворонок В. В. Знаки української етнокультури: Словник-довідник.— К.: Довіра, 2006.— 703 с.
30. Жайворонок В. В. Велика чи мала літера? Словник-довідник.— К.: Наук. думка, 2004.— 203 с.
31. Загнітко А. П., Данилюк І. Г., Ситар Г. В., Щукіна І. А. Словник українських прийменників: Сучасна українська мова.— Донецьк: ТОВ ВКФ «БАО», 2007.— 416 с.
32. Зализняк А. А. Грамматический словарь русского языка: Словоизменение.— М.: Рус. яз., 1980.— 880 с.
33. Золотова Г. А. Синтаксический словарь. Репертуар элементарных единиц русского синтаксиса.— М.: Наука, 1988.— 439 с.
34. Іванця В. О., Підгорський В. С., Юргелайтіс Н. Г. та ін. Словник термінів у мікробіології.— К.: Наук. думка, 2006.— 200 с.

35. Інверсійний словник української мови / Відп. ред. С. П. Бевзенко. — К. : Наук. думка, 1985. — 812 с.
36. Їсакевич Г. П., Калюжна В. М., Паламарчук О. В. та ін. Українсько-російський словник. — К. : Наук. думка, 2004. — 984 с.
37. Калашиник В. С., Колоїз Ж. В. Словник фразеологічних антонімів української мови. — К. : Довіра, 2001. — 284 с.
38. Караванський С. Й. Практичний словник синонімів української мови. — К. : Довіра, 2004. — 275 с.
39. Караванський С. Й. Російсько-український словник складної лексики. — К. : Академія, 1998. — 709 с.
40. Караулов Ю. Н., Черкасова Г. А., Уфимцева Н. В. и др. Русский ассоциативный словарь : В 2 т. — М. : АСТ-Астрель, 2002.
41. Карачун В. Я. Орфографічний словник наукових і технічних термінів. — К. : Криниця, 1999. — 528 с.
42. Карачун В. Я., Карачун О. А., Прохур Ю. З., Гульчук Г. Г. Російсько-український словник з інформатики та обчислювальної техніки. — К. : Рось, 1994. — 362 с.
43. Карачун В. Я., Кремень Т. В., Прохур Ю. З. Російсько-українсько-англійський економіко-математичний словник. — К. : Грамота, 2008. — 440 с.
44. Карпіловська С. А. Кореневий гніздовий словник української мови : Гнізда слів з вершинами — омографічними коренями. — К. : Укр. енцикл. ім. М. П. Бажана, 2002. — 912 с.
45. Короткий тлумачний словник української мови / За ред. Д. Г. Гринчишина. — 2-ге вид., перероб. і доп. — К. : Рад. шк., 1988. — 320 с.
46. Культура мови на щодень / Н. Я. Дзюбишина-Мельник, Н. С. Дужик, С. Я. Єрмоленко та ін. — К. : Довіра, 2000. — 170 с.
47. Куньч З. Словарь русско-украинских языковых омонимов. — К. : Академія, 1997. — 400 с.
48. Лингвистический энциклопедический словарь / Под ред. В. Н. Ярцева. — М. : Науч. изд-во «Большая Рос. энцикл.», 2002. — 972 с.
49. Лучик А. А. Російсько-український та українсько-російський словник еквівалентів слова / Відп. ред. М. М. Пещак. — К. : Довіра, 2003. — 495 с.
50. Лучик А. А. Словник еквівалентів слова української мови. — К. : Вид. дім «Киево-Могилянська академія», 2008. — 174 с.
51. Морковкин В. В., Луцкая Н. М., Богачева Г. Ф. Объяснительный словарь русского языка. — 2-е изд., испр. — М. : АСТ-Астрель, 2003. — 421 с.
52. Морозов С. М., Шкаранута Л. М. Словник іншомовних слів. — К. : Наук. думка, 2000. — 683 с.
53. Мюллер В. К. Великий англо-український словник. — К. : Арія, 2008. — 784 с.
54. Новый объяснительный словарь синонимов русского языка : Проспект / Под ред. Ю. Д. Апресяна. — М. : Рус. словари, 1995. — 560 с.
55. Нові слова та значення : Словник / Л. В. Туровська, Л. М. Василькова. — К. : Довіра, 2008. — 271 с.

56. *Ожегов С. И.* Словарь русского языка / Под ред. Н. Ю. Шведовой. — 18-е изд., стер. — М.: Рус. яз., 1987. — 797 с.
57. *Ожегов С. И., Шведова Н. Ю.* Толковый словарь русского языка. — 4-е изд., доп. — М.: ООО «ИТИ ТЕХНОЛОГИИ», 2003. — 944 с.
58. Орфоепічний словник української мови : В 2 т. / За ред. М. М. Пешак, В. М. Русанівського. — К.: Довіра, 2001–2003.
59. Орфоепічний словник української мови / В. М. Русанівський, В. В. Чу-мак, Г. М. Ярун. — К.; Ірпінь: ВТФ «Перун», 2006. — 208 с.
60. Повний словник антонімів української мови / В. С. Калашник, Ж. В. Колоіз. — К.: Довіра, 2006. — 859 с.
61. Політехнічний російсько-український словник. — К.; Ірпінь: ВТФ «Перун», 2000. — 512 с.
62. *Полюга Л. М.* Словник антонімів української мови / За ред. Л. С. Паламарчука. — 2-ге вид., доп. і випр. — К.: Довіра, 2001. — 276 с.
63. *Полюга Л. М.* Словник українських морфем. — 3-тє вид. — К.: Довіра, 2009. — 554 с.
64. *Полюга Л. М.* Словник українських синонімів та антонімів. — К.: Довіра, 2007. — 575 с.
65. *Полюга Л. М.* Словник фразеологічних антонімів української мови. — К.: Довіра, 2006. — 859 с.
66. Практичний російсько-український словник найуживаніших слів і словосполучень / За ред. Ф. А. Непійводи. — К.: Основи, 2000. — 256 с.
67. Російсько-український словник / За ред. Г. П. Півторака. — К.: Довіра, 2002. — 632 с.
68. Російсько-український словник / За ред. В. В. Жайворонка. — К.: Абрис, 2003. — 1424 с.
69. Російсько-український словник: В 3 т. / Редкол.: І. К. Білодід (голова) та ін. — 3-тє вид. — К.: УРЕ, 1987–1988.
70. Російсько-український словник з радіотехніки, радіоелектроніки та радіофізики / За ред. В. С. Калашника та М. Т. Черпака. — К.: Наук. думка, 2006. — 680 с.
71. Російсько-український словник наукової термінології: Біологія. Хімія. Медицина. — К.: Наук. думка, 1996. — 664 с.
72. Російсько-український словник наукової термінології: Математика. Фізика. Техніка. Науки про Землю та Космос. — К.: Наук. думка, 1998. — 892 с.
73. Російсько-український словник наукової термінології: Суспільні науки. — К.: Наук. думка, 1994. — 600 с.
74. Російсько-український словник-довідник. — 2-е вид. / За ред. В. М. Брїцина. — К.: Довіра, 2008. — 942 с.
75. Русский семантический словарь. Толковый словарь, систематизированный по классам слов и значений / Под общ. ред. Н. Ю. Шведовой. — М.: Изд. центр «Азбуковник», 1998. — Т. I. — 832 с.
76. *Саврук М. П.* Українсько-англійський науково-технічний словник. — К.: Наук. думка, 2008. — 912 с.

77. Словарь русского языка : В 4 т. / Под ред. А. П. Евгеньевой. — 2-е изд., испр. и доп. — М. : Рус. яз., 1981–1984.
78. Словарь современного русского литературного языка : В 17 т. — М. : Изд-во АН СССР, 1948–1965.
79. Словарь української мови : В 4 т. / Упоряд., з дод. власного матеріалу Б. Грінченко. — К. : Наук. думка, 1996–1997.
80. Словник античної міфології / Відп. ред. А. О. Білецький. — К. : Наук. думка, 1985. — 235 с.
81. Словник іншомовних слів / За ред. О. С. Мельничука. — 2-ге вид., випр. і доп. — К. : УРЕ, 1985. — 966 с.
82. Словник іншомовних слів: 23 000 слів та термінологічних словосполучень / Уклад. Л. О. Пустовіт, О. І. Скопненко, Г. М. Сюта та ін. — К. : Довіра, 2000. — 1018 с.
83. Сучасний словник іншомовних слів / Уклад. О. І. Скопненко, Т. В. Цимбалюк. — К. : Довіра, 2006. — 789 с.
84. Словник синонімів української мови : В 2 т. / А. А. Бурячок, Г. М. Гнатюк, С. І. Головашук. — К. : Наук. думка, 2006.
85. Словник української мови : В 11 т. / Гол. ред. І. К. Білодід. — К. : Наук. думка, 1970–1980.
86. Словник української мови : У 20 т. — Т. 1 : А–Б / Гол. наук. ред. В. М. Русанівський. — К. : Наук. думка, 2010. — 911 с.
87. Словник фразеологізмів української мови / Відп. ред. В. О. Винник. — К. : Наук. думка, 2003. — 1098 с.
88. Українська мова. Енциклопедія / За ред. В. М. Русанівського. — К. : Укр. енцикл. ім. М. П. Бажана, 2000. — 750 с.
89. Український орфографічний словник / За ред. В. Г. Скляренка — 9-те вид., перероб. і доп. — К. : Довіра, 2009. — 1011 с.
90. Український орфографічний словник / За ред. В. М. Русанівського. — К. : Дніпро, 2006. — 940 с.
91. Українсько-російський словник : В 6 т. / Гол. ред. І. М. Кириченко. — К. : Вид-во АН УРСР, 1953–1963.
92. Українсько-російський словник наукової термінології / За зап. ред. Л. О. Симоненко. — К. ; Ірпінь : ВТФ «Перун», 2004. — 416 с.
93. Фразеологічний словник української мови : В 2 кн. / За ред. Л. С. Паламарчука. — К. : Наук. думка, 1993.
94. Христианство. Энциклопедический словарь : В 2 т. — М. : Науч. изд-во «Большая Рос. энцикл.», 1993–1995.
95. Частотний словник сучасної української художньої прози : В 2 т. — К. : Наук. думка, 1981.
96. Шевченко В. М. Словник-довідник з релігієзнавства. — К. : Наук. думка, 2004. — 559 с.
97. Языкознание. Большой энциклопедический словарь / Гл. ред. В. Н. Ярцева. — 2-е изд. — М. : Науч. изд-во «Большая Рос. энцикл.», 1988. — 685 с.

98. Яценко І. Т. Морфемний аналіз : Словник-довідник : В 2 т. — К. : Вища шк., 1980–1981.
99. A Dictionary of Accounting.— Oxford University Press, 2005.— 408 с.
100. A Dictionary of Law.— Second edition published.— Glasgow : Harper Collins Publishers, 2001.— 416 p.
101. Barnhart C. A. The Facts On File Student's dictionary of american english.— N.Y. : Facts On File, 2008.— 737 p.
102. Concise Oxford Russian Dictionary.— Revised Edition.— М. : Весь мир, Изд. дом ИНФРА-М, 2005.— 1007 с.
103. Oxford Primary Dictionary : Толковый словарь английского языка / Р. Аллен.— М. : АСТ-Астрель, 2007.— 568 с.
104. The Oxford Dictionary of the English Language: Оксфордский толковый словарь английского языка / Дж. Хокинс.— Oxford University Press, М. : АСТ-Астрель, 2007.— 828 с.
105. The Oxford Illustrated English Dictionary : Оксфордский толковый иллюстрированный словарь английского языка / Р. Аллен.— М. : АСТ-Астрель, 2003.— 560 с.

Наукові праці

106. Агрикола Э. Микро-, мезо- и макроструктура как содержательная основа словаря // Вопр. языкознания.— 1984.— № 2.— С. 72–87.
107. Анализ метаязыка словаря с помощью ЭВМ / Ю. Н. Караулов, В. И. Молчанов, В. А. Афанасьев, Н. В. Михалев.— М. : Наука, 1982.— 94 с.
108. Апресян Ю. Д. Избранные труды : В 2 т.— М. : Шк. «Языки рус. культуры», 1995.
109. Апресян Ю. Д. О языке толкований и семантических примитивах // Известия АН.— Сер. лит. и яз.— 1994.— Т. 53.— № 4.— С. 27–41.
110. Апресян Ю. Д. Образ человека по данным языка: попытка системного описания // Вопр. языкознания.— 1995.— № 1.— С. 38.
111. Арбатский Д. И. Толкования значений слов. Семантические определения.— Ижевск : Удмуртия, 1977.— 100 с.
112. Ахманова О. С. Очерки по общей и русской лексикологии.— М. : Учпедгиз, 1957.— 295 с.
113. Бабкин А. М. Идеоматика и грамматика в словаре // Современная русская лексикография.— Л. : Наука, 1981.— С. 5–43.
114. Балли Ш. Общая лингвистика и вопросы французского языка / Ред., вступ. ст. и примеч. Р. А. Будагова.— М. : Изд-во иностр. лит., 1971.— 416 с.
115. Бацевич Ф. С., Космеда Т. А. Очерки по функциональной лексикологии.— Львов : Свит, 1997.— 392 с.
116. Безпояско О. К., Городенська К. Г., Русанівський В. М. Грамматика української мови. Морфологія : Підр.— К. : Либідь, 1993.— 336 с.
117. Бенвенист Э. Общая лингвистика / Под ред. Ю. С. Степанова.— М. : Прогресс, 1974.— 445 с.

118. *Березан С. Г.* Компонентный анализ и словарь (соотношение между семантическим и лексикографическим метаязыками) // Семантика языка и текста.— Тез. между. науч. конф. 28–30 октября 1987 года.— Кировоград, 1987.— С. 6–7.
119. *Бехтерев В. М.* Объективная психология.— М.: Наука, 1991.— 480 с.
120. *Біблія, або книги Святого Письма Старого і Нового Заповіту із мови давньоєврейської і грецької на українську наново перекладена.*— К.: Укр. Біблійне тов-во, 2002.— 1165 с.
121. *Білоноженко В. М., Гнатюк І. С.* Функціонування та лексикографічна розробка українських фразеологізмів.— К.: Наук. думка, 1989.— 155 с.
122. *Бланар В.* Лексикология и лексикография // *Вопр. языкознания.*— 1985.— № 3.— С. 77–82.
123. *Блумфилд Л.* Язык / Пер. с англ.— М.: Прогресс, 1968.— 607 с.
124. *Будагов Р. А.* Толковые словари в национальной культуре народов.— М.: Изд-во Моск. ун-та, 1989.— 151 с.
125. *Васильев Л. М.* Современная лингвистическая семантика.— М.: Высш. шк., 1990.— 176 с.
126. *Везжица А.* Введение // Семантические примитивы / Семиотика / Под ред. Ю. С. Степанова.— М., 1983.— С. 225–252.
127. *Везжица А.* Семантические универсалии и «примитивное мышление» // Язык. Культура. Познание.— М.: Рус. словари, 1996.— 411 с.
128. *Выготский Л. С.* Мышление и речь.— М.: Лабиринт.— 1996.— 416 с.
129. *Виноградов В. В.* Лексикология и лексикография: Избр. труды.— М.: Наука, 1977.— 312 с.
130. *Виноградов В. В.* Русский язык (Грамматическое учение о слове): Учеб. пособие для вузов / Отв. ред. Г. А. Золотова.— 3-е изд., испр.— М.: Высш. шк., 1986.— 640 с.
131. *Винокур Г. О.* Толковый словарь и языковое употребление // *Вопр. языкознания.*— 1986.— № 4.— С. 16–26.
132. *Вихованець І. Р.* Граматика української мови. Синтаксис: Підр.— К.: Либідь, 1993.— 368 с.
133. *Волкова И. Н.* Моделирование определений в терминологических стандартах // Современные проблемы русской терминологии.— М.: Наука, 1986.— С. 140–162.
134. *Гак В. Г.* Языковые преобразования.— М.: Шк. «Языки рус. культуры», 1998.— 763 с.
135. *Гейзенберг В.* Физика и философия. Часть и целое / Пер. с нем.— М.: Наука, 1989.— С. 191–196.
136. *Гнатюк М. Г.* Граматичний аспект у лексикографії // *Мовознавство.*— 1967.— № 3.— С. 30–38.
137. *Головацук С. І.* Перекладні словники і принципи їх укладання.— К.: Наук. думка, 1976.— 247 с.
138. *Головка Г. Г., Широков В. А.* Теорема Левенгейма-Сколема как формальный коррелят лексикографического эффекта в информационных системах // *Горизонти прикладної лінгвістики та лінгвістичних технологій (MegaLing'2006):*

Доп. другої міжнар. наук. конф., 20–27 вересня, 2006, Україна, Крим, Партевіт.— С. 201–202.

139. *Горецький П. Й.* Історія української лексикографії.— К.: Вид-во АН УРСР, 1963.— 243 с.

140. *Городецкий Б. Ю.* Проблемы и методы современной лексикографии // Новое в зарубежной лингвистике.— М.: Прогресс, 1983.— Вып. 14.— С. 5–23.

141. *Городецкий Б. Ю., Раскин В. В.* Методы семантического исследования ограниченного подязыка.— М.: Изд-во Моск. ун-та, 1971.— 414 с.

142. Граматика української мови / За ред. І. Р. Вихованця, К. Г. Городенської, А. П. Грищенка.— К.: Рад. шк., 1982.— 208 с.

143. Граматика української мови: Морфологія / За ред. О. К. Безпояско, К. Г. Городенської, В. М. Русанівського.— К.: Либідь, 1993.— 335 с.

144. *Грицев С. В.* Введение в терминологическую лексикографию.— М.: Изд-во Моск. ун-та, 1986.— 106 с.

145. *Грищенко П. Ю.* Моделирование системы диалектной лексики.— К.: Наук. думка, 1984.— 227 с.

146. *Гумбольдт В.* Избранные труды по языкознанию.— М.: Прогресс, 1984.— 397 с.

147. *Девкин В. Д.* О неродившихся немецких и русских словарях // Вопр. языкознания.— 2001.— № 1.— С. 85–97.

148. *Денисов П. Н.* Лексика русского языка и принципы ее описания.— М.: Рус. яз., 1980.— 253 с.

149. *Денисов П. Н.* Системность и связанность в лексике и система словарей // Проблематика определений терминов в словарях разных типов.— Л.: Наука, 1976.— С. 63–73.

150. *Дирак П. А. М.* Принципы квантовой механики.— 2-е изд.— М.: Наука, 1979.— 408 с.

151. *Дорошевський В.* Про глуначення значень слів у словниках // Мовознавство.— 1970.— № 5.— С. 10–20.

152. *Дьячков М. В.* Миноритарные языки в полиэтнических (многонациональных) государствах.— М.: ИНЦО, 1996.— 116 с.

153. *Евгеньева А. П.* Определение в толковых словарях // Проблема толкования слова в филологических словарях.— Рига: Изд-во АН ЛатССР, 1963.— С. 20.

154. *Есперсен О.* Философия грамматики / Пер. с англ. В. В. Пассека и С. П. Сафроновой, Под ред. и с предисл. Б. А. Ильица.— М.: Изд-во иностр. лит., 1958.— 404 с.

155. *Жовтобрюх М. А., Кулик Б. М.* Курс сучасної української літературної мови: У 2 т.— К.: Вища шк., 1972.— 402 с.

156. *Журавлев В. Ф.* По поводу определений в формальной логике, в сборниках рекомендуемых терминов, терминологических стандартах, толковых словарях и энциклопедиях // Проблематика определений терминов в словарях разных типов.— Л.: Наука, 1976.— С. 114–121.

157. *Загнітко А. П.* Морфологічні категорії іменника в синтагматичі і парадигматичі.— К.: УМКВО, 1989.— 63 с.

158. *Загнітко А. П.* Система і структура морфологічних категорій сучасної української мови : Проблеми теорії : Дис. ... д-ра філол. наук / АН України.— К., 1992.— 453 с.
159. *Загнітко А. П.* Теоретична граматика української мови. Морфологія.— Донецьк : Вид-во Донецьк. держ. ун-ту, 1996.— 435 с.
160. *Заде Л.* Понятие лингвистической переменной и его применение к принятию приближенных решений.— М. : Мир, 1976.— 167 с.
161. *Зализняк А. А.* Грамматический словарь русского языка : Словоизменение.— М. : Рус. яз., 1980.— 880 с.
162. *Золотова Г. А.* Синтаксический словарь. Репертуар элементарных единиц русского синтаксиса.— М. : Наука, 1988.— 439 с.
163. Інформаційні матеріали Наукової ради з проблем «Закономірності розвитку національних мов у зв'язку з розвитком соціалістичних націй» / Відп. ред. М. А. Жовтобрюх.— К. : Наук. думка, 1974.— Вип. 16.— 80 с.
164. *Караулов Ю. Н.* Лингвистическое конструирование и тезаурус литературного языка.— М. : Наука, 1981.— 367 с.
165. *Караулов Ю. Н.* Общая и русская идеография.— М. : Наука, 1976.— 355 с.
166. *Караулов Ю. Н., Молчанов В. И., Афанасьев В. А., Михалев Н. В.* Анализ метаязыка словаря с помощью ЭВМ.— М. : Наука, 1982.— 94 с.
167. *Карри Г.* Некоторые логические аспекты грамматической структуры / Пер. с англ. В. В. Лазарева // Новое в лингвистике.— М. : Прогресс, 1965.— Вып. 4.— С. 97-116.
168. *Касарес Х.* Введение в современную лексикографию.— М. : Изд-во иностран. лит., 1958.— 354 с.
169. *Кацнельсон С. Д.* Содержание слова, значение и обозначение.— Л. : Наука, 1965.— 110 с.
170. *Клименко Н. Ф., Пещак М. М., Савченко І. Ф.* Формалізовані основи семантичної класифікації лексики / Відп. ред. М. М. Пещак.— К. : Наук. думка, 1982.— 250 с.
171. *Кобозева И. М.* Лингвистическая семантика.— М. : КомКнига, 2007.— 352 с.
172. *Кодухов В. И.* Контекст как лингвистическое понятие // Языковые единицы и контекст.— Л. : Просвещение, 1973.— С. 7-32.
173. *Колмогоров А. М.* Три подхода к определению понятия «количества информации» // Теория информации и теория алгоритмов.— М. : Наука, 1987.— 304 с.
174. *Колишанский Г. В.* Контекстная семантика.— М. : Наука, 1980.— 149 с.
175. *Комарова З. И.* Проблема толкования специальных слов в словарях русского языка советского периода : Учеб. пособие.— Свердловск : Изд-во Уральского гос. ун-та, 1979.— 91 с.
176. *Коннолли Т., Бегг К., Страчан А.* Базы данных : проектирование, реализация и сопровождение. Теория и практика.— 2-е изд.— М. : Изд. дом «Вильямс», 2000.— 1120 с.

177. *Котелова Н. З.* Значение слова и его сочетаемость (к формализации в языкознании). — Л. : Наука, 1975. — 164 с.
178. *Кочерган М. П.* Слово і контекст (Лексична сполучуваність і значення слова). — Львів : Вища шк., Вид-во Львів. ун-ту, 1980. — 183 с.
179. *Кузнецов А. М.* От компонентного анализа к компонентному синтезу. — М. : Наука, 1986. — 126 с.
180. *Кузнецова Э. В.* Лексикология русского языка : Учеб. пособие для филол. фак. ун-тов. — 2-е изд. испр. и доп. — М. : Высш. шк., 1989. — 216 с.
181. *Кузнецова Э. В.* Ступенчатая идентификация как средство описания семантических связей слов // *Вопр. металингвистики.* — Л., 1973. — С. 84-94.
182. *Лексическая основа русского языка : Комплексный учеб. словарь / Под ред. В. В. Морковкина.* — М. : Рус. яз., 1984. — 1167 с.
183. *Лисиченко Л. А.* Лексикологія сучасної української мови : Семантична структура. — Харків : Вища шк., 1977. — 144 с.
184. *Лосев А. Ф.* Знак. Символ. Миф : Труды по языкознанию. — М. : Изд-во Моск. ун-та, 1982. — 479 с.
185. *Лучик А. А.* Семантика прислівникових еквівалентів слова української і російської мов. — К. : Довіра, 2001. — 217 с.
186. *Максимов В. И.* Структура и членение слова. — Л. : Изд-во Ленингр. ун-та, 1977. — 147 с.
187. *Манако В. В., Якименко К. М.* Автоматизоване тестування тексту фундаментальної академічної лексикографічної системи «Словник української мови» // *Горизонти прикладної лінгвістики та лінгвістичні технології (MegaLing'2006) : Доп. Міжнар. конф. — Партевіт, 20-27 вер. 2006. — Сімферополь : «ДиАйПи». — 2006. — С. 150.*
188. *Манако В. В., Якименко К. М.* Технологические аспекты создания текста первого тома фундаментальной академической лексикографической системы «Словарь украинского языка» // *Прикладна лінгвістика та лінгвістичні технології (MegaLing'2007) : Зб. наук. праць.* — К. : Довіра. — 2008. — С. 233-242.
189. *Медникова Е. М.* Значение слова и методы его описания. — М. : Высш. шк., 1974. — 202 с.
190. *Межжерина Г. В.* Структурна організація семантичних одиниць (поле – лексико-семантична група – слово) // *Актуальні проблеми української лінгвістики. теорія і практика.* — К., 2002. — Вип. V. — С. 113-126.
191. *Мельничук О. С.* Словник української мови в 11 томах // *Вісник АН УРСР.* — 1984. — № 3. — С. 100-103.
192. *Мельчук И. А.* Опыт теории лингвистических моделей «смысл – текст» : Семантика, синтаксис. — М. : Шк. «Языки рус. культуры», 1999. — 345 с.
193. *Мельчук И. А.* Строение языковых знаков и возможные формально-смысловые отношения между ними // *Изв. АН СССР. — Сер. лит. и яз.* — 1968. — Т. 27. — Вып. 5. — С. 426-438.
194. *Методи структурного дослідження мови / Відп. ред. В. С. Перебийніс, Л. О. Кадомцева.* — К. : Наук. думка, 1968. — 187 с.

195. *Москаленко А. А.* Нарис історії української лексикографії.— К.: Рад. шк., 1961.— 162 с.
196. *Назарова І. С.* Лексикографічна розробка українського дієслова (про деякі дієслівні категорії): Дис. ... канд. філол. наук / АН УРСР.— К., 1969.— 285 с.
197. Национальная специфика языка и ее отражение в нормативном словаре: Сб. ст. / Отв. ред. Ю. Н. Караулов.— М.: Наука, 1988.— 172 с.
198. *Нечитайло О. І.* Синоніми в лексикографії.— К.: Наук. думка, 1987.— 132 с.
199. *Німчук В. В.* Староукраїнська лексикографія в її зв'язках з російською та білоруською.— К.: Наук. думка, 1980.— 304 с.
200. Новое в русской лексике. Словарные материалы – 77 / Под ред. Н. З. Котеловой.— М.: Рус. яз., 1980.— 176 с.
201. Новые слова и словари новых слов: Сб. ст. / Отв. ред. Н. З. Котелова.— Л.: Наука, 1982.— 184 с.
202. *Ожегов С. И.* Лексикология. Лексикография. Культура речи: Учеб. пособие для вузов.— М.: Высш. шк., 1974.— 352 с.
203. *Ожегов С. И.* О трех типах толковых словарей.— М.: Наука, 1983.— 215 с.
204. *Озерова Н. Г.* Лексическая и грамматическая семантика существительного.— К.: Наук. думка, 1990.— 192 с.
205. *Падучева Е. В.* Высказывание и его соотносительность с действительностью.— М.: Наука, 1985.— 271 с.
206. *Паламарчук Л. С.* Лексико-семантичний розвиток мови // Мовознавство.— 1992.— № 4.— С. 3–8.
207. *Паламарчук Л. С.* Новий академічний словник // Мовознавство.— 1980.— № 5.— С. 3–9.
208. *Паламарчук Л. С.* Нові номінації як один з аспектів збагачення словника // Слово і труд: Зб. наук. праць.— К.: Наук. думка, 1976.— 210 с.
209. *Паламарчук Л. С.* Про принципи добору лексичного інвентаря до загальнономовного словника // Мовознавство.— 1973.— № 3.— С. 3–11.
210. *Паламарчук Л. С.* Тлумачний словник української мови в колі слов'янських словників цього типу.— К.: Наук. думка, 1973.— С. 14.
211. *Паламарчук Л. С.* Українська радянська лексикографія: питання історії, теорії і практики.— К.: Наук. думка, 1978.— 203 с.
212. *Паламарчук Л. С., Юрчук Л. А.* Засади створення Словника української мови // Мовознавство.— 1967.— № 3.— С. 19–29.
213. *Пецак М. М.* Комунікативний синтаксис.— К.: Довіра, 2000.— 150 с.
214. *Пецак М. М.* Нариси з комп'ютерної лінгвістики.— Ужгород: Закарпаття, 1999.— 199 с.
215. *Пецак М. М.* Стан і перспективи комп'ютерної лексикографії в Україні // Мовознавство.— 1996.— № 4/5.— С. 8–11.
216. *Пецак М. М., Климченко Н. Ф., Карпиловская Е. А. и др.* Украинский семантический словарь. Проспект (машинный формат базы данных и принципы его автоматического составления).— К.: Наук. думка, 1990.— 264 с.

217. *Пещак М. М., Клименко Н. Ф., Ярун Г. М., Карпіловська Є. А.* Лексична семантика в системі «людина — машина». — К. : Наук. думка, 1986. — 282 с.
218. *Погрібна О. О.* Комплексні постулати представлення граматичної семантики іменників української мови у тлумачному словнику // *Мовні і концептуальні картини світу.* — К. : ВПУ «Київський університет», 2004. — С. 176–182.
219. *Погрібна О. О.* Лексична та граматична семантика іменників у «Словнику української мови» (назви знарядь праці та споруд) : Дис. ... канд. філол. наук / НАН України. Ін-т мовознавства ім. О. О. Потебні. — К., 2005. — 245 с.
220. *Погрібна О. О., Чумак В. В.* Формалізація структури граматичної семантики лівої частини словникової частини іменників у 11-томному тлумачному словнику // *Прикладна лінгвістика та лінгвістичні технології (Megaling-2009)* : 36. наук. праць. — К. : Довіра, 2009. — С. 347–357.
221. *Погрібна О. О., Чумак В. В., Широков В. А., Шевченко І. В.* Лінгвістична класифікація українського іменника у світлі теорії лексикографічних систем // *Мовознавство*. — 2004. — № 5/6. — С. 62–82.
222. *Потебня А. А.* Из записок по русской грамматике. — М. : Учпедгиз, 1958. — Т. 1–2. — 536 с.
223. *Пуаза Б.* Курс теории моделей / Пер. с фр. (электронный учебник). — Алматы : Казахский нац. ун-т, 2001. — 474 с.
224. *Рабулець О. Г.* Інтегровані лексикографічні системи : Автореф. дис. ... канд. техн. наук: 05.13.06 / НАН України; Національна бібліотека України ім. В. І. Вернадського. — К., 2002. — 18 с.
225. *Рабулець О. Г., Сухарина Н. М., Широков В. А., Якименко К. М.* Дієслово в лексикографічній системі. — К. : Довіра, 2004. — 259 с.
226. *Раевская М. М.* Язык в ментальном пространстве: к проблеме постижения национальной логики мышления // *Вестник МГУ.* — Сер. 19 : Лингвистика и межкультурная коммуникация. — 2006. — № 1. — С. 27–41.
227. Развитие лексики современного русского языка / Под ред. Е. А. Земской, Д. Н. Шмелева. — М. : Наука, 1965. — 136 с.
228. Распределённые системы. Принципы и парадигмы / Э. Тененбаум, М. ванн Стеен. — СПб. : Питер, 2003. — 877 с.
229. *Ревзина О. Г.* Контекст // *Русский язык : Энциклопедия* / Под ред. Ю. Н. Караулова — 2-е изд., перераб. и доп. — М. : Науч. изд-во «Большая Рос. энцикл.» ; Дрофа, 1998. — С. 198.
230. *Рогожникова Р. П.* Словарь эквивалентов слова: наречные, служебные, модальные единства. — М. : Рус. яз., 1991. — 254 с.
231. *Розина Р. И.* Семантическое развитие слова в русском литературном языке и в современном сленге: глагол. — М. : Изд. центр «Азбуковник», 2005. — 302 с.
232. *Роменская В. Ф.* О классификационной схеме понятия «тип словаря» в информационном тезаурусе // *Структурная и прикладная лингвистика.* — Л., 1978. — Вып. 1. — С. 181–187.
233. *Руднев В. П.* Словарь культуры XX века : Ключевые понятия и тексты. — М. : Аграф, 1999. — 384 с.

234. *Русанівський В. М.* Дієслово — рух, дія, образ.— К. : Рад. шк., 1977.— 111 с.
235. *Русанівський В. М.* Закономірності розвитку значення слова // *Мовознавство*.— 1981.— № 4.— С. 27—36.
236. *Русанівський В. М.* Структура лексичної і граматичної семантики.— К. : Наук. думка, 1988.— 237 с.
237. *Русанівський В. М.* Структура українського дієслова.— К. : Наук. думка, 1971.— 315 с.
238. *Русанівський В. М., Гриценко П. Ю.* Здобутки і проблеми світового слов'янознавства: До підсумків IX Міжнародного з'їзду славистів // *Вісник АН УРСР*— 1984.— № 3.— С. 77—88.
239. *Русанівський В. М., Єрмоленко С. Я.* Життя слова.— К. : Вища шк. при КДУ, 1978.— 191 с.
240. *Русанівський В. М., Тараненко О. О., Широков В. А.* Теоретико-лінгвістичні засади та інформаційно-комп'ютерне забезпечення україномовних лінгвістичних інтелектуальних систем // *Мовознавство*.— 1996.— № 4/5.— С. 3—7.
241. *Русанівський В. М., Широков В. А.* Інформаційно-лінгвістичні основи сучасної тлумачної лексикографії // *Мовознавство*.— 2002.— № 6.— С. 3—8.
242. *Русская грамматика* : В 2 т. / Под ред. Н. Ю. Шведовой.— М. : Наука, 1980.
243. *Скороходько Э. Ф.* Семантические сети и автоматическая обработка текста.— К.: Наук. думка, 1983.— 218 с.
244. Слово і фразеологізм у словнику : Зб. наук. праць / Відп. ред. Л. С. Паламарчук.— К. : Наук. думка, 1980.— 250 с.
245. *Смирницкий А. И.* Лексическое и грамматическое в слове // *Вопр. грамматического строя*.— М. : Изд-во АН СССР, 1955.— С. 11—53.
246. *Сніжко М. Д.* Ідеографічний тезаурус як модель лексико-семантичної системи (за наслідками автоматизованого аналізу українських іменників) // *Мовознавство*.— 1995.— № 6.— С. 28—35.
247. *Современность и словари* : Сб. ст. / Ред. Ф. П. Филин, Ф. П. Сороколетов.— Л. : Наука, 1978.— 181 с.
248. *Соколов О. М.* Основы имплицитной морфологии русского языка.— М. : Изд-во РУДН, 1997.— 203 с.
249. *Соколовская Ж. П.* Проблемы системного описания лексической семантики.— К. : Наук. думка, 1990.— 184 с.
250. *Солнцев В. И.* Язык как системно-структурное образование.— 2-е изд.— М. : Недра, 1977.— 341 с.
251. *Соломоник А.* Семиотика и лингвистика.— М. : Молодая гвардия, 1995.— 352 с.
252. Сопоставительная стилистика русского и украинского языков / Отв. ред. В. И. Кононенко.— К. : Вища шк., 1980.— 204 с.
253. *Соссюр Ф. де.* Курс общей лингвистики / Пер. со 2-го фр. изд. А. М. Сухотина.— М. : Логос, 1999.— 296 с.

254. *Стернин И. А.* Проблемы анализа структуры значения слова.— Воронеж : Изд-во Воронеж. гос. ун-та, 1979.— С. 76.
255. *Стратулат Н. В.* Семантична неологізація як спосіб збагачення словникового складу української мови // Мовознавство.— 2007.— № 3.— С. 69–77.
256. *Сухарина Н. М.* Граматична та лексична семантика українського дієслова в лексикографічній системі : Дис... канд. філол. наук / НАН України.— К., 2003.— 234 с.
257. Сучасна українська літературна мова : В 5 кн. / За заг. ред. І. К. Білодіда.— К. : Наук. думка, 1969–1973.— Кн. 1 : Фонетика.— 1969.— 436 с.; Кн. 2 : Морфологія.— 1969.— 583 с.; Кн. 3 : Синтаксис.— 1972.— 515 с.; Кн. 4 : Лексика і фразеологія.— 1973.— 439 с.; Кн. 5 : Стилїстика.— 1973.— 588 с.
258. *Тараненко А. А.* Языковая семантика в ее динамических аспектах (основные семантические процессы).— К. : Наук. думка, 1989.— 256 с.
259. *Тараненко О. О.* Новый словарь украинської мови: Концепція і принципи укладання словника.— К. ; Кам'янець-Подільський, 1996.— 172 с.
260. Толковый словарь русского языка конца XX в. Языковые изменения / Под ред. Г. Н. Складарской.— СПб. : Фолио-Пресс, 1998.— Введение.— С. 7–36.
261. *Томіленко Л. М.* Морфологічний спосіб творення нових іменників-термінів у Словнику української мови в 20 томах // Мовознавство.— 2009.— № 2.— С. 69–78.
262. *Ульман Дж.* Основы систем баз данных.— М. : Финансы и статистика, 1983.— 334 с.
263. *Успенский В. А.* К определению падежа по А. Н. Колмогорову. Бюллетень Объединения по проблемам машинного перевода.— № 5.— М. : [И МГПИИЯ], 1957.— С. 11–18.
264. *Успенский В. А.* Теорема Гёделя о неполноте.— М. : Наука, 1982.— 110 с.
265. *Уфимцева А. А.* Типы словесных знаков.— М. : Наука, 1974.— 206 с.
266. *Чейф У.* Значение и структура языка.— М. : Прогресс, 1975.— 432 с.
267. *Шафф А.* Введение в семантику / Пер. с польского Н. Комлева, В. Конновой.— М. : Изд-во иностр. лит., 1963.— 376 с.
268. *Шведова Н. Ю.* Лексическая система и ее отражение в толковом словаре // Русистика сегодня. Язык: система и ее функционирование.— М. : Наука, 1988.— 224 с.
269. *Шведова Н. Ю.* Однотомный толковый словарь: специфика жанра и работы // Русский язык : Проблемы художественной речи. Лексикология и лексикография. (Виноградовские чтения. 9–10).— М., 1981.— С. 171.
270. *Шевченко І. В.* Моделі та алгоритмічно-програмне забезпечення лексикографічних систем : Автореф. дис. ... канд. техн. наук: 05.13.06 / НАН України.— К., 2000.— 19 с.
271. *Шевченко Л. Л.* Лексика Святого Письма в 20-томному Словнику української мови // Українсько-македонський наук. зб.— К., 2009.— Вип. 4.— С. 417–424.

272. *Шевченко Л. Л., Томіленко Л. М.* Мотиваційні основи та базові моделі семантичного процесу термінологізації (на матеріалі Словника української мови в 20 томах) // *Мовознавство*.— 2010.— № 1.— С. 76–85.
273. *Шенк Р., Бирнбаум Л., Мей Дж.* К интеграции семантики и прагматики // *Новое в зарубежной лингвистике : Вып. 24. Компьютерная лингвистика*.— М., 1988.— С. 33.
274. *Шилейко А. В., Кочнев В. Ф., Химушин Ф. Ф.* Введение в информационную теорию систем.— М. : Радио и связь, 1985.— 278 с.
275. *Широков В. А.* Елементи лексикографії.— К. : Довіра, 2005.— 304 с.
276. *Широков В. А.* Інформаційна теорія лексикографічних систем.— К. : Довіра, 1998.— 331 с.
277. *Широков В. А.* Інформаційна теорія та системотехнічні засади комп'ютерної лексикографії : Дис. ... д-ра техн. наук / НАН України.— К., 1999.— 384 с.
278. *Широков В. А.* Нова класифікація українських дієслів // *Актуальні проблеми української лінгвістики : Теорія і практика : Зб. праць*.— 2003.— Вип. 8.— С. 113–125.
279. *Широков В. А.* О скрытой симметрии лексикографической системы Словаря украинского языка // *Прикладна лінгвістика та лінгвістичні технології (MegaLing'2007) : Зб. наук. праць / За ред. В. А. Широкова*.— К. : Довіра, 2008.— С. 403–424.
280. *Широков В. А.* Семантические состояния языковых единиц // *Прикладная лингвистика в поиске новых путей (MegaLing'2005) : Тр. междунар. конф. / Отв. ред. В. П. Захаров, С. С. Дикарева*.— Спб. : Изд-во «Осипов», 2005.— 180 с.— С. 147–162.
281. *Широков В. А.* Семантичні стани мовних одиниць та їх застосування в когнітивній лексикографії // *Мовознавство*.— 2005.— № 3–4.
282. *Широков В. А.* Феноменологія лексикографічних систем.— К. : Наук. думка, 2004.— 327 с.
283. *Широков В. А., Бугаков О. В., Грязнухіна Т. О. та ін.* Корпусна лінгвістика.— К. : Довіра, 2005.— 471 с.
284. *Широков В. А., Воронько М. П., Костышин А. М.* Структура и информационная модель толкового Словаря украинского языка // *Проблеми створення машинних фондів мов : Тр. науч. конф.*— К., 1991.— С. 74–76.
285. *Широков В. А., Манако В. В.* Організація ресурсів Національної словникової бази // *Мовознавство*.— 2001.— № 5.— С. 3–13.
286. *Широков В. А., Манако В. В.* Технологічний комплекс зі створення фундаментальної академічної лексикографічної системи «Словник української мови» // *VII Міжнар. наук.-практ. конф. «Інформаційні ресурси наук.-техн. інформації: проблеми створення і використання»*.— К. : УкрІНТЕІ.— 2000.— С. 121–124.
287. *Широков В. А., Пещак М. М.* Структурна модель реєстрової частини Словника української мови // *Національна архівна інформаційна система «Архівна та рукописна Україніка і комп'ютеризація архівної справи» : Зб. наук. праць*.— Вип. 1 : «Інформатизація архівної справи в Україні: сучасний стан та перспективи».— К., 1996.— С. 154–174.

288. Широков В. А., Рабулець О. Г. Формалізація у галузі лінгвістики // Актуальні проблеми української лінгвістики: теорія і практика : Зб. праць.— К., 2002.— Вип. 5.— С. 3–27.

289. Широков В. А., Рабулець О. Г., Костишин О. М. та ін. Технологічні основи сучасної тлумачної лексикографії // Мовознавство.— 2002.— № 6.— С. 49–86.

290. Широков В. А., Рабулець О. Г., Шевченко І. В., Якименко К. М. Інтегрована лексикографічна система «Словники України», в. 4.1 (електронне видання).

291. Широков В. А., Сухарина Н. М., Якименко К. М. Лексикографічна система як джерело лінгвістичних класифікацій // Вісн. Київ. лінгв. ун-ту.— 2003.— Т. 6.— № 2.— С. 43–53.

292. Широков В. А., Шевченко І. В., Рабулець О. Г. та ін. Словники України — інтегрована лексикографічна система : Парадигма. Транскрипція. Фразеологія. Синонімія. Антонімія.— К., 2001.— 1 електрон. опт. диск (cdr).— (Словники України).

293. Широков К. В. Іменна словозміна у сучасній турецькій мові.— К. : Довіра, 2009.— 318 с.

294. Широков К. В., Широков В. А. Застосування формалізму нечітких множин для визначення граматичних станів турецьких слів // Мовознавство.— 2005.— № 5.— С. 51–56.

295. Шмелев Д. Н. Проблемы семантического анализа лексики (на материале рус. языка).— М. : Наука, 1973.— 280 с.

296. Шрам А. Н. Структурные типы лексических значений слова // Филологические науки.— 1981.— № 2.— С. 60.

297. Щерба Л. В. Избранные работы по языкознанию и фонетике.— Т. 1.— Л. : ЛГУ, 1958.— 182 с.

298. Щерба Л. В. Опыт общей теории лексикографии // Избр. работы по языкознанию и фонетике.— Л. : Изд-во Ленингр. гос. ун-та, 1958.— Т. 1.— С. 54–91.

299. Щерба Л. В. Языковая система и речевая деятельность.— Л. : Наука, 1974.— 428 с.

300. Шур Г. С. Теория поля в лингвистике.— М. : Наука, 1974.— 255 с.

301. Юрчук Л. А. Ілюстративний матеріал як засіб семантичної характеристики слова в тлумачному словнику // Лексикологія та лексикографія : Респуб. міжвід. зб.— К. : Наук. думка, 1969.— Вип. 3.— С. 132–140.

302. Юрчук Л. А. Про типи тлумачень дієслів з префіксом до- у сучасній українській мові // Лексикологія та лексикографія : Респуб. міжвід. зб.— К. : Наук. думка, 1966.— Вип. 2.— С. 114–122.

303. Юрчук Л. А. Теоретичні засади реєстру фразеологічного словника української мови // Мовознавство.— 1983.— № 5.— С. 23–32.

304. Язык и наука конца XX века : Сб. ст. / Под ред. Ю. С. Степанова.— М. : Ин-т языкознания РАН, 1995.— 432 с.

305. *Якименко К. М.* Системотехніка та технологія лексикографічних систем семантичного типу : Автореф. дис. ... канд. техн. наук: 05.13.06 / НАН України.— К., 2006.— 18 с.
306. ANSI/X3/SPARK DBMS study group interim report. // FDT-Bull. ACM SIGMOD.— 1975.— V. 7.— № 2.— 140 p.
307. *Bugakov O.* Definitions of Prepositions, Conjunctions and Particles in the Explanatory Dictionaries // Proceedings of MONDILEX Fourth Open Workshop : Warsaw, 2009.— P. 194–197.
308. *Chomsky N.* Language in a Psychological Setting // *Sophia Linguistica* (Toyo), 1987.— P. 21.
309. CODASYL DBTG 1971 [CODASYL Data Base Task Group April 71 Report. ACM New York, 1971].
310. Extensible Markup Language (XML) 1.0 [Електронний ресурс]: W3C Recommendation 04 February 2004.— Third Edition.— Режим доступу: <http://www.w3.org/TR/2004/REC-xml-20040204>.
311. *Morris Ch.* Signs, language and behaviour.— N. Y. : Moss, R.N., 1947.— 164 p.
312. *Rabulets A.* System Engineering Principles of Virtual Linguistic Laboratories // Organisation and Development of Digital Lexical Resources. Proceedings of Mondilex Second Open Workshop.— Kyiv, 2009.— P. 18–23.
313. *Shyrovok V.* Experience in Creating a National Dictionary Depository of Ukraine and its Use in Conceptual Modelling of Networking of Centres for High-quality Research in Slavic Lexicography and their Digital Resources // Organisation and Development of Digital Lexical Resources. Proceedings of Mondilex Second Open Workshop.— Kyiv, 2009.— P. 5–9.
314. *Shyrovok V.* Theory of Lexicographic systems : Part 1 // Metalanguage and Encoding Scheme Design for Digital Lexicography. Proceedings of Mondilex Third Open Workshop.— Bratislava, 2009.— P. 151–167.
315. *Shyrovok V.* Theory of Lexicographic systems : Part 2 // Representing Semantics in Digital Lexicography. Proceedings of Mondilex Fourth Open Workshop.— Warszawa, 2009.— P. 89–105.
316. *Shyrovok V.* Theory of Lexicographic systems : Part 3 // Research Infrastructure in Digital Lexicography. Proceedings of Mondilex Fifth Open Workshop.— Ljubljana, 2009.— P. 98–119.
317. *Tsichritzis D.* and *Klug A.* (eds.) The ANSI/X3/SPARK Framework.— AFIPS Press, Nontvale, N. J., 1978.
318. W3C WSDL 2.0 0 [Електронний ресурс]: Core Specifications («Мова опису Web-сервісів (мова WSDL), версія 2.0»)— Режим доступу: www.w3.org/TR/wsdl20.
319. *Zgusta L.* Manual of Lexicography.— Prague : Publishing house of the Czechoslovak Academy of Sciences, 1971.— 357 p.

Вступ	3
Розділ 1. Тлумачна лексикографія на загальнолінгвістичному тлі	13
1.1. Означення, функції та загальна структура словника	13
1.2. Лексичне значення слова та його лексикографічне представлення	19
1.3. Концепція станів мовних одиниць	31
1.4. Формалізований підхід до визначення семантичних станів слів	37
Розділ 2. Теорія лексикографічних систем	45
2.1. Лексикографічний ефект в інформаційних системах	45
2.2. Структура і архітектура лексикографічних систем	57
2.3. Лексикографічні структури і словники	73
2.4. Лексикографічні середовища	81
2.5. Інтегровані Л-системи та методика їх побудови	82
Розділ 3. Системно-лінгвістичні аспекти тлумачного Словника української мови	87
3.1. Теорія Л-системи тлумачного Словника української мови	87
3.2. Прихована симетрія українського дієслова	92
3.3. Іменник. Структура та класифікація	110
3.4. Мікроструктура інтерпретаційних частин словникових статей СУМа	136
3.5. Формалізований підхід до аналізу словникових дефініцій дієслів у СУМі	143
3.6. Формалізований підхід до аналізу словникових дефініцій іменників у СУМі	150

Розділ 4. Головні засади укладання СУМ-20	160
4.1. Основні принципи лексикографування	160
4.2. Лексичний склад Словника. Критерії добору лексики та формування реєстру	166
4.3. Орфографія СУМ-20	170
Розділ 5. Подання мовного матеріалу в СУМ-20	176
5.1. Структура словникової статті	176
5.2. Будова лівої частини словникової статті. Лексикографічний опис реєстрового ряду	177
5.3. Будова правої частини словникової статті. Тлумачення реєстрових одиниць	179
5.4. Лексикографічна характеристика частин мови	184
5.5. Подання словосполучень	191
5.6. Лексика і фразеологія Святого Письма	200
5.7. Нова загальноживана і термінологічна лексика	206
5.8. Стилїстична характеристика реєстрових слів	215
5.9. Текстово-ілюстративний матеріал	220
5.10. Поліграфічне оформлення СУМ-20	224
5.11. Індекс СУМ-20	232
Розділ 6. Технологічні аспекти проекту «СУМ-20»	236
6.1. Парсинг СУМ-11	236
6.2. Структура і функції лексикографічної бази даних та інструментального комплексу «СУМ»	240
6.3. Віртуальна лексикографічна лабораторія «СУМ-20»	249
6.4. Лінгвістичний корпус	257
6.5. Організація лексикографічного процесу при укладанні СУМ-20	266
Література	277

Наукове видання

Колектив авторів

**ЛІНГВІСТИЧНІ
ТА ТЕХНОЛОГІЧНІ ОСНОВИ
ТЛУМАЧНОЇ ЛЕКСИКОГРАФІЇ**

Відповідальний редактор *В. А. Широков*

Редактор *Н. В. Щербакова*

Художній редактор *В. П. Мариняк*

Коректор *В. П. Зуб*

Комп'ютерна верстка *М. Ю. Кригіна, І. В. Остапової*

Формат 60x84 ¹/₁₆. Умовн. друк. арк. 17,2.
Наклад 300 прим. Замовлення 9-194.

Видавництво «Довіра»

вул. Кіквідзе, 2/34, Київ-103, 01103

Свідоцтво про внесення до Державного реєстру видавців
серія ДК № 669 від 14.11.2001 р.

ЗАТ «Книга», вул. Артема, 25, Київ-53, 04053

Свідоцтво про внесення до державного реєстру виготівників
серія ДК № 2323 від 25.10.2005 р.

Лінгвістичні та технологічні основи тлумачної лексикографії /
Л59 В. А. Широков, В. М. Білоноженко, О. В. Бугаков та ін. — К. :
Довіра, 2010. — 295 с. : іл. — Бібліогр. : С. 277-293.
ISBN 978-966-507-283-6

У монографії викладено лінгвістичні та системотехнічні принципи створення великих тлумачних лексиконів – основних об'єктів національної лексикографії. Сформульовано теоретичні підходи до побудови великих словників на основі розроблених авторами теорії лексикографічних систем і теорії семантичних станів. Теоретичні підходи докладно проілюстровано численними прикладами з практики укладання фундаментальних академічних словників української мови в 11 та 20 томах.

ББК 81.2УКР-4