

КОМП'ЮТЕРНА ЛІНГВІСТИКА

*Ганна Ситар
Донецький національний університет*

СТРУКТУРА БАЗИ ДАНИХ “СИНТАКСИЧНІ ФРАЗЕОЛОГІЗМИ В УКРАЇНСЬКІЙ МОВІ”

У статті запропоновано параметри опису синтаксичних фразеологізмів. Висвітлено два етапи створення бази даних “Синтаксичні фразеологізми в українській мові”, розглянуто її структуру.

Ключові слова: база даних, модель речення, синтаксичний фразеологізм, фразеологізоване речення, українська мова.

Сучасний етап розвитку лінгвістики характеризується видозміною традиційних підходів до вивчення мови й мовлення та застосуванням нових методів аналізу мовних явищ і мовленнєвих процесів. У межах структурної, а пізніше й комп'ютерної лінгвістики, мовну інформацію розглядають як особливий різновид інформації (докладно про ознаки мовної інформації див.: [Белецкий : 11–18; Карпіловська : 14–15]), що зумовлює залучення до її дослідження методів інших галузей знань, які вивчають процеси одержання, репрезентації, зберігання й оброблення інформації, – інформатики, кібернетики і програмування.

В арсеналі комп'ютерної лінгвістики провідне місце посідають бази даних (далі БД) як найпоширеніший спосіб організації мовних даних. За К. Дж. Дейтом, база даних – комп'ютеризована система зберігання записів, що містить великий обсяг інформації й надає її для потреб користувача [Дейт].

БД можуть бути застосовані в різних галузях мовознавства, у випадках, коли виникає потреба проаналізувати значний за обсягом масив мовних одиниць, здійснити багатовимірну класифікацію, визначити корелятивність / некорелятивність кількох параметрів даних [Рахилина; Ситар, Милостна та ін.].

Дані, об'єднані в межах БД, мають єдиний формат, властивості внутрішньої інтерпретованості, структурованості, пов'язаності

ності та ін. [Поспелов]. Саме тому БД видається найефективнішим інструментом для компактної репрезентації сукупності параметрів синтаксичних фразеологізмів, зручності та швидкості оброблення потрібних даних.

Теоретичним підґрунтям створення БД “Синтаксичні фразеологізми в українській мові” став аналіз поглядів лінгвістів на місце цих одиниць у системі мови, їхні структурні, семантичні й прагматичні властивості [Балобанова; Величко; Всеволодова, Лим Су Ен; Личук 2001; Личук 2001; РГ-80; Хохлина; Шведова; Шмелев 2006; Шмелев 2006 та ін.]. Услід за Н. Шведовою, А. Величко, М. Всеволодовою та іншими вченими синтаксичний фразеологізм (або фразеологізоване речення) розуміємо як особливий тип речення, у якому компоненти пов’язані ідіоматично, синтаксичні зв’язки й прями лексичні значення слів послаблені, наприклад:

– *Там ти його навряд чи впізнаєш, – застерігає Ліду батько.*

– *Кирило Петрович там до невпізнання опроститься, сиву бороду відпустить, уподібнившись тим східним мудрецам, що йшли колись у гори й роками непорушно сиділи серед скель, заглиблені в самоспоглядання... <...> **Чим не перспектива?** Тренуватиме бджіл, шкрябатиме мемуари, сидячи, ніби якийсь маг бородатий, серед степу широкого на Україні милій... (О. Гончар);*

Отець Харитін знов прийняв нігтя й почав читати, але Онисія Степанівна вже зовсім не слухала. Отець Харитін примітив те й спитав:

– *Чи ви слухаєте?*

– *Та слухаю, слухаю. Не питай мене та читай далі.*

– *Коли ви щось думаєте, – сказав отець Харитін, придивляючись до розумного лиця своєї жінки.*

– *Читай-бо, читай! Я думаю й слухаю заразом.*

– *А як звали Варвариного батька?*

– *А хіба ж ти читав про батька?*

– *Авжеж читав, тільки ви не слухали.*

– *А правда, правда! То був якийсь ігумен.*

– **Який там ігумен!** Діоскор – так звали Варвариного батька.

– Та хто ж такий був ігумен? Чи який Варварин родич, чи що? – спитала вона.

– Не ігумен-бо, а Ігемон! Це той, що мучив святу Варвару.

– Ото проклятий! В якому ж монастирі він був ігемоном? І чого ж то ченці мучили святу Варвару?

– **Які там ченці!** Ігемон – то так звався губернатор, чи становий би то по-нашому, чи що; от так, як в Богуславі отой становий, що в Чайках намальований у пеклі (І. Нечуй-Левицький).

Фразеологізоване речення складається з двох частин – змінної й постійної, в останній порядок компонентів є фіксованим, наприклад:

– **Оце так зустріч!** – не знаходить Дмитро слів, щоб висловити свою радість, і не чує, як з-під сплетення засніжених вій виступають дві прозорі краплини.

– Неждана і хороша, – усміхається Кошовий і захоплено оглядає Горлиці з усіх сторін (М. Стельмах);

Оце так крос! Українець Сергій Лебідь переміг на 15-му чемпіонаті Європи з кросу на 10 км у Брюсселі – 30.49! Це вже восьме “золото” нашого легендарного 33-річного кросмена, який цього разу завдяки фірмовому фінішному спурту випередив найближчого переслідувача Мо Фараха з Великої Британії на 500 м! (Високий замок. – 16.12.2008).

Мета цієї статті – охарактеризувати структуру БД “Синтаксичні фразеологізми в українській мові”, зокрема:

1) висвітлити основні етапи створення БД “Синтаксичні фразеологізми в українській мові”;

2) виділити параметри опису синтаксичних фразеологізмів;

3) розглянути структуру БД, навести приклади зроблених записів і заповнених полів.

Проектування БД здійснювалося у два етапи (докладно про етапи створення БД див.: [Карпіловська : 34–35]). Перший – інфологічний – етап включав відбір і структурування відповідних

мовних і/або лінгвістичних даних, тобто виділення моделей синтаксичних фразеологізмів в українській мові та встановлення їхніх класифікаційних ознак. На сьогодні наш реєстр нараховує близько 100 моделей синтаксичних фразеологізмів (з урахуванням їхніх варіантів), дібраних із творів українських письменників ХІХ–ХХІ ст., з преси останніх п'ятнадцяти років, а також зафіксованих у мережі Інтернет і в усному мовленні.

Аналіз зібраного фактичного матеріалу й досвіду лексикографічного опису російських традиційних (нефразеологізованих) [Русские...], нечленованих [Меликян 2001; Меликян 2005] і фразеологізованих [Балобанова; Меликян 2000; Меликян 2005] речень засвідчують необхідність багатоаспектного аналізу синтаксичних фразеологізмів як особливого типу мовних одиниць, що перебувають на межі синтаксису і фразеології [Ситар 2009; Ситар 2011; Ситар 2011 та ін.]. На наш погляд, достатніми для коректного їхнього опису є такі параметри: 1) структурна схема речення; 2) типова семантика речення; 3) лексичні обмеження моделі; 4) тип речення за будовою (просте, складне – з подальшим розмежуванням підтипів); 5) тип речення за частиномовним статусом незмінного компонента; 6) ступінь злитості компонентів; 7) варіанти структурної схеми речення; 8) додаткові семантичні відтінки; 9) семантичний тип (семантична група); 10) можливість / неможливість уведення поширювачів; 11) парадигма речення; 12) образні моделі речень; 13) прагматична функція синтаксичного фразеологізму; 14) тип прагматичного зв'язку з іншими репліками діалогу; 15) прагматичний статус мовця (тип відношення мовця до ситуації); 16) семантико-парадигматичні властивості моделі (синонімічні й омонімічні речення).

Датологічний етап проектування БД передбачав вибір системи керування базою даних (СКБД). На сьогодні відомі кілька типів БД: ієрархічні, мережеві й реляційні [Поспелов та ін.]. БД “Синтаксичні фразеологізми в українській мові” побудована за реляційною моделлю. На відміну від ієрархічної або мережевої, реляційна БД містить “інформацію про той чи інший мовний об'єкт як сукупність його атрибутів, властивостей, характеристик, ознак” [Карпіловська : 35–36]. Створювана БД стано-

вить сукупність структурованих даних про синтаксичні фразеологізми української мови, тобто включає моделі синтаксичних фразеологізмів різного типу, їх характеристики.

Як зазначає К. Рахіліна, лексична БД реляційного типу є таблицею із двома входами – лексемами та їхніми ознаками, які приписуються мовним одиницям у відповідних рядках таблиці [Рахіліна : 108]. У нашому випадку входами таблиці є не лексеми, а моделі речень і параметри пропонованої класифікації.

Для створення БД була обрана програма СКБД Microsoft Access, оскільки вона є найпоширенішою реляційною СКБД у Microsoft Windows і забезпечує зберігання, сортування й пошук даних [Канноли, Бег : 283]. Під час роботи з нею користувач має справу з кількома об'єктами, основними з них є таблиці, форми й запити. Таблиці, що становлять основу БД, складаються з полів і рядків (записів). Кожній моделі синтаксичного фразеологізму в режимі таблиці надано індивідуальний номер.

Для зручності роботи з БД була сконструйована так звана форма зі вкладками, що об'єднують кілька полів, елементами керування й навігації. Структура БД корелює із запропонованими параметрами описуваних мовних одиниць. Основним полем у режимі форми обрано поле “Структурна схема речення”. Інші поля, що відповідають виділеним властивостям синтаксичних фразеологізмів, згруповані для зручної роботи користувача в сім вкладок.

1. Вкладка “*Структура*” відображає такі властивості, як: “Тип речення за будовою”, “Частиномовний статус незмінного компонента”, “Варіанти структурної схеми”, “Наявність поширювачів” (рис. 1).

2. Вкладка “*Семантика*” охоплює поля “Типова семантика речення”, “Додаткові семантичні відтінки”, “Семантичний тип”, “Ступінь злитості компонентів”, “Лексичні обмеження”, “Образні моделі” (рис. 2). Зазначимо, що останні два поля для деяких синтаксичних фразеологізмів залишаються незаповненими, оскільки існують синтаксичні фразеологізми, що не мають обмежень на лексичне заповнення змінюваної позиції (наприклад, модель *Чим не N₁* (докладніше про це див.: [Ситар 2011]), а образні моделі реалізуються в межах не всіх проаналізованих речень.

Синтаксичні фразеологізми в українській мові 1.02

Структурна схема речення:

Структура **Семантика** Синтаксична парадигматика Прагматика Семантико-парадигматичні властивості Словникові відомості Приклади

Тип речення за будовою:

Частини мови статус незмінного компонента:

Варіанти структурної схеми:

Наявність поширювачів:

а) А що стосується Євро-2012... Чим не привід для гордоше?... (Дзеркало тижня. – 11-18.11.2007 (№ 42)); б) У Дніпропетровську в обласному дербі зійдуться «Дніпро» й «Кривбас», чим не можливість прорекламувати свої тренерські таланти для наставника криворізьців Юрія Максимова? (Високий замок. – 22.09.2010).

а) Та в Юнаві налаштовані оптимістично: Китай надав Україні гарантії ядерної безпеки, у нас немає проблемних питань, держави постійно підтримують одна одну в міжнародних організаціях, товаробіг в останні роки постійно зростає, чим не стратегічний партнер? (Дзеркало тижня. – 18-24.10.2003 (№ 40)); б) От зіграли ви Берію колись. А зараз сидите і на телевізор кося поглядаєте, а там Савакашвілі зі співвітчизниками воює, чим не іронія історії? (Дзеркало тижня. – 17-23.11.2007 (№ 44)); в) Ринок нині перебуває в тій торш, коли через дватри місяці м'яса бракуватиме, а ціни зростатимуть, чим не привід заробити грошей? (Дзеркало тижня. –

Рис. 1. Структурні параметри моделі Чим не N₁ Corp_f (режим форми)

Синтаксичні фразеологізми в українській мові 1.02

Структурна схема речення:

Структура **Семантика** Синтаксична парадигматика Прагматика Семантико-парадигматичні властивості Словникові відомості Приклади

Типова семантика речення:

Додаткові семантичні відтінки:

Семантичний тип:

Ступінь злитості компонентів:

Лексичні обмеження:

Образні моделі:

а) Нічого, звісно ж, не виходило, проте він [Янгол] знову й знову вперто стрибав на кожній наступній програвці. Люто дивився у небо і знову поривався туди. І знову тягла до себе земля. Стрибок – приземлення, стрибок – приземлення, чим тобі не культурний відпочинок психа? (Тетяна Вінокурова-Садиченко. Жарт. із життя психів);

Господар! Справжній господар. Ремесло в руках має. Життя мозолем своїм добився. Земля вироблена. За таким жінка добряче проживе. І собою вдає – чим не орел? Так!

Рис. 2. Семантичні параметри моделі Чим не N₁ Corp_f (режим форми)

3. Вкладка “Синтаксична парадигматика” включає відомості про компонентний склад синтаксичної парадигми описуваної моделі речення – граматичні модифікації часу, способу, фазові, модальні, заперечні, авторизаційні та інші перетворення базової моделі речення. Парадигма синтаксичних фразеологізмів, звичайно, є неповною, але виявлення обмежень на модифікації для кожної окремої моделі видається значущим для опису її властивостей (рис. 3).

Синтаксичні фразеологізми в українській мові 1.02

Структурна схема речення:

Наявність синтаксичної парадигми:

Модифікації часу:

Модифікації способу:

Фазові модифікації:

Модальні модифікації:

Заперечні модифікації:

Авторизаційні модифікації:

Інші перетворення:

Рис. 3. Синтаксична парадигма моделі *Чим не N₁ Corp_f* (режим форми)

4. Вкладка “Прагматика” поєднує три поля – “Прагматична функція”, “Прагматичний статус мовця” і “Тип прагматичного зв’язку”.



5. Вкладка “Семантико-парадигматичні властивості” відображає синонімічні зв’язки моделей синтаксичних фразеологізмів з іншими фразеологізованими моделями й традиційними реченнями; синтаксичні омоніми розмежовано на нефразеологізовані і фразеологізовані речення, побудовані за однією структурною схемою (рис. 4).

6. Вкладка “Словникові відомості” об’єднує дані, які наведені в авторитетних словниках. Оскільки на сьогодні в україністиці моделі речень ще не зазнали лексикографічного опрацювання

(пор. словники російських нефразеологізованих і фразеологізованих речень [Русские...; Меликян 2000; Меликян 2005]), мова йде про відомості, почерпнуті із фразеологічних або тлумачних словників, у яких зафіксовані лише окремі синтаксичні фразеологізми. Зазначимо, що для моделі синтаксичного фразеологізму *Чим не N₁ Cop_f* ця вкладка залишилася порожньою.

Синтаксичні фразеологізми в українській мові 1.02

Структурна схема речення:

Структура	Семантика	Синтаксична парадигматика	Прагматика	Семантико-парадигматичні властивості	Словникові відомості	Приклади
Синтаксичні синоніми:		<input type="text" value="Оце (так) N1 Copf! Оце N1 так N1! Це N1Copf! Це справний N1Copf!"/>				
Синтаксичні омоніми:		<input type="text" value="відсутні"/>				

Рис. 4. Семантико-парадигматичні властивості моделі *Чим не N₁ Cop_f* (режим форми)

7. У вкладці “Приклади” подано контексти вживання певної моделі синтаксичних фразеологізмів переважно в художньому, публіцистичному й розмовному стилях. Тут подаються приклади основного варіанта моделі, а випадки неелементарної структури, ускладненої семантики й утворення модифікацій ілюструються прикладами в межах відповідних вкладок.

Отже, БД “Синтаксичні фразеологізми в українській мові” спроектована за реляційною моделлю. На сьогодні вона включає дані про основні структурні, семантичні, прагматичні й парадигматичні властивості виявлених моделей синтаксичних фразеологізмів.

Створена версія БД не є остаточним варіантом. Наступні етапи роботи передбачають не тільки поповнення новими моделями, а й розв’язання низки теоретичних і прикладних питань, зокрема проблеми виділення додаткових диференційних ознак

синтаксичних фразеологізмів, розмежування в їх межах членованих і нечленованих одиниць, установлення функційного навантаження варіантності пунктуаційного оформлення, що відкриває перспективи подальшого теоретичного вивчення закономірностей фразеологізації речень і застосування отриманих результатів для укладання словника синтаксичних фразеологізмів української мови.

1. *Балобанова Л. А.* Семантико-прагматический потенциал синтаксических фразеологизмов и их лексикографическое представление в словаре учебного типа: Дис. ... канд. пед. наук. – М., 2004.
2. *Белецкий А. А.* Семиотический аспект языковой системы // Структурная и математическая лингвистика. – К., 1979. – Вып. 7. – С.11–18.
3. *Величко А. В.* Синтаксическая фразеология для русских и иностранцев. – М., 1996.
4. *Всеволодова М. В., Лим Су Ён.* Принципы лингвистического описания синтаксических фразеологизмов: На материале синтаксических фразеологизмов со значением оценки. – М., 2002.
5. *Дейт К. Дж.* Введение в системы баз данных. – 6-е изд. – К., 1998.
6. *Канноли Т., Бег К.* Базы данных. Проектирование, реализация и сопровождение. Теория и практика. – М., 2003.
7. *Карпіловська Є. А.* Вступ до прикладної лінгвістики: комп'ютерна лінгвістика. – Донецьк, 2006.
8. *Личук М. І.* Ступені фразеологізації речень: Автореф. дис. ... канд. філол. наук. – К., 2001.
9. *Личук М. І.* Ступені фразеологізації речень. – Чернівці, 2001.
10. *Меликян В. Ю.* Словарь фразеосинтаксических схем русского языка. – Ростов-на-Дону, 2000.
11. *Меликян В. Ю.* Эмоционально-экспрессивные обороты живой речи: Словарь. – М., 2001.
12. *Меликян В. Ю., Краснов В. А.* Словарь коммуникем русского языка со значением "волеизъявления". – Ростов-на-Дону, 2005.
13. *Меликян В. Ю., Остапенко А.И.* Словарь фразеологизированных сложноподчиненных предложений русского языка. – Ростов-на-Дону, 2005.
14. *Поспелов Д. А.* Данные и знания // Искусственный интеллект: В 3-х кн. – Кн. 2. Модели и методы: Справочник / Под ред. Д. А. Поспелова. – М., 1990. – С. 7–14.
15. *Рахилина Е. В.* О лексических базах данных // Вопросы языкознания. – 1994. – № 4. – С. 107–113.
16. *Русская грамматика: В 2-х т. – Т. 2: Синтаксис / Под ред. Н. Ю. Шведовой.* – М., 1980.
17. *Русские глагольные предложения: Экспериментальный синтаксический словарь / Под общ. ред. Л. Г. Бабенко.* – М., 2002.
18. *Ситар Ганна.* Синтаксичні фразеологізми з модальним значенням в українській мові // *Ukrainistika – minulost, přítomnost, budoucnost II. Sborník vědeckých prací.* – Brno, 2009. – S. 270–276.
19. *Ситар Ганна.* Синтаксичні фразеологізми, побудовані за моделлю *Чим не N₁*, в українській мові // Лінгвістичні студії: зб. наук. праць. – Донецьк, 2011. – Вип. 22. – С. 208–215.
20. *Ситар Ганна.* Теоретичні основи словника синтаксичних фразеологізмів: структура словникової статті // Лінгвістичні студії: зб. наук. праць. – Донецьк,

2011. – Вип. 23. – С. 288–292. 21. *Ситар Ганна, Милостна Олена*. База даних як інструмент лінгвістичного дослідження (на прикладі аналізу лексичних конверсивів) // Лінгвістичні студії: зб. наук. праць. Вип. 16. – Донецьк, 2008. – С. 385–392. 22. *Хохлина М. Л.* Синтаксическая идиоматика русского языка в лингвопрагматическом аспекте: На материале устойчивых моделей простого предложения: Автореф. дис. ... канд. филол. наук. – Астрахань, 2006. 23. *Шведова Н. Ю.* О некоторых типах фразеологизированных конструкций в строе русской разговорной речи // Вопросы языкознания. – 1958. – № 2. – С. 93–100. 24. *Шмелев Д. Н.* Синтаксическая членимость высказывания в современном русском языке. – Изд. 2-е, стереотипное. – М., 2006. 25. *Шмелев Д. Н.* Современный русский язык. Лексика. – М., 2006.

В статье предложены параметры описания синтаксических фразеологизмов. Освещены два этапа создания базы данных “Синтаксические фразеологизмы в украинском языке”, рассмотрена ее структура.

Ключевые слова: база данных, модель предложения, синтаксический фразеологизм, фразеологизированное предложение, украинский язык.

The paper presents the parameters of describing the syntactic idioms. The two stages of creating a database “Syntactic idioms in the Ukrainian language” are elucidated, and its structure is regarded.

Key words: database, sentence model, syntactic idiom, sentence with phraseological structure, the Ukrainian language.