

Електронний словник синтаксем: організація мікроструктури та перспективи застосування

С. Б. Фокін

Київський національний університет імені Тараса Шевченка
Corresponding author. E-mail: sergiyborysovych@ukr.net

Paper received 23.11.18; Revised 26.11.18; Accepted for publication 27.11.18.

<https://doi.org/10.31174/SEND-HS2018-184VI30-04>

Анотація. Сучасна лексикографія характеризується виникненням принципово нових об'єктів, з-поміж яких виокремлюються граматичні словники. Хоча граматичні словники можна укладати відповідно до різних моделей та з різними цілями, особливо корисними для численних практичних і науково-теоретичних застосувань є словники синтаксем. Оптимальна організація такого словника (за семасіологічним і ономасіологічним принципами) забезпечить можливість різного роду пошуку в усіх можливих частинах макро- та мікроструктури, а також укладання двомовного словника синтаксем на його основі.

Ключові слова: синтаксема, лексикографія, ономасіологія, електронний словник, граматична семантика, функціональний синтаксис, переклад.

Вступ. Сучасна комп'ютерна лексикографія характеризується бурхливим, якщо не сказати експоненційним, ритмом розвитку. Ще два десятиліття тому суттєвим досягненням вважалося переведення паперового словника у цифровий формат, а сьогодні вже зовсім не дивує, що, окрім великого розмаїття електронних словників, укладених людиною, існує ціла низка лексикографічних ресурсів, автоматично згенерованих комп'ютером. Переважно це стосується укладання конкордансів, частотних словників, зворотних і інверсійних словників, а також інших об'єктів лінгвістичного конструювання.

Розробники комп'ютерних ресурсів у першу чергу активно користуються тими можливостями, які елементарно реалізувати шляхом простої формалізації. У той самий час існують об'єктивні потреби у розробці словників, принципово нових не лише в формальному, а й у змістовому відношенні. З сімдесятих-вісімдесятих років ХХ століття починають з'являтися доволі екзотичні “нелексічні” словники у паперовому форматі. Оскільки процес створення цих словників не піддається формальному опису й не може бути доручений комп'ютерам, ними не рясніють сучасні комп'ютерні ресурси, доступні онлайн, тому вони заслуговують на детальну згадку.

Огляд публікацій за темою. Мова йде про словники, основні складові одиниці яких виходять за межі лексики на інші мовні, мовленнєві і мовно-культурні рівні: орфоепічні словники, словники епітетів, фразеологічні словники, словники цитат, типових помилок (в українській традиції – так звані “словник культури мови”) тощо. Однак особливо несподіваним і новітнім видається укладання граматичних словників, позаяк сполучення “граматична лексикографія” звучить як оксюморон, хоча, як справедливо стверджує К.Г. Городенська, це повноцінний окремий напрямок лексикографії [4], у рамках якого можна назвати такі розробки: “Грамматичний словник української мови. Сполучники” [3] (упорядкований К.Г. Городенською), “Грамматичний словник української літературної мови: словозміна” (укладений В.І. Критською [7]). І.Р. Вихованець працює над розробкою “Словника українських прийменників. Сучасна українська літературна мова” [4, с. 32].

Коротко окреслюючи історію питання, зауважимо,

що її витоки слід шукати не власне у лексикографії (і тим паче, комп'ютерній), а радше у функціональній граматиці. У цілому в основі функціональної граматики лежать ідеї В. Матезіуса, О.В. Бондарка, І.Р. Вихованця, Г.А. Золотової, В.П. Даниленка та інших мовознавців. Ідеї цих учених розвиваються, на превеликий жаль, відокремлено від комп'ютерної лінгвістики. Хоча в ній самій вже відбувся відчутний поступ, і сучасні технології дозволяють здійснювати напівавтоматичну частиномовну розмітку тексту, а деякі лінгвісти підходять впритул також до лексико-семантичного анування, ще зарано говорити про розмітку корпусів тексту у термінах функціональної граматики і поняттєвих категорій. І це не тому, що такі категорії як аспектуальність, темпоральність, модальність, стан, персональність, кількісність, часова локалізованість, локативність, посесивність, буттєвість та інші недоступні для комп'ютерної техніки, а через те, що знання функціональної граматики, які можна отримати тільки на спеціальних факультетах вишів, не проникають у кола розробників комп'ютерних програм та й, відверто кажучи, не надто поширені й з-поміж мовознавців.

Серед словників, які ґрунтуються безпосередньо на ідеях функціонального синтаксису, окрім згаданих праць К.Г. Городенської, І.Р. Вихованця, назвемо також “Іспансько-російсько-український словник граматичних лексем”, упорядкований Е.В. Шевкун [8]. Спільним для словника Е.В. Шевкун і К.Г. Городенської є наявність значної кількості сполучників, сполучникових зв'язкових виразів, що супроводжуються описом їхніх функцій. Шляхом семантичного, синтаксичного, морфологічного аналізу І.Р. Вихованець, довів, що провідна функція службових частин мови, на противагу повнозначним, – синтаксична, тому вчений їх умовно називає аналітичними синтаксичними морфемами [1, с. 19]. А це означає, що саме такі службові мовленнєві одиниці як сполучники в словниках доречно розглядати саме у складі синтаксем. Відомо, що в двомовних перекладних словниках службові частини мови, такі як сполучники, прийменники, частки, важко піддаються встановленню еквівалентних відношень з лексемами цільової мови, що призводить до непомірного розростання словникових статей і обтяжливих пояснень. Не надто легким для розуміння сполучників

виявляються й пояснення високого рівня абстракції у тлумачних словниках. Однак опис їхніх функцій, ілюстративні приклади реалізації цих функцій у складі словосполучень, що являють собою синтаксеми, значно полегшують сприйняття знань.

Таким чином, функціональний підхід до граматики поступово потрапляє до словників, і це вже слід визнати колосальним поступом у розвитку лексикографії. Наступним важливим кроком, на нашу думку, має стати розвиток цього напрямку у двомовну перекладну лексикографію. Адаже за умов розуміння функції слова чи сполучення слів можливо віднаходити належний функціональний відповідник в іншій мові на чіткій методологічній основі. Особливо цінно для користувача те, що еквівалентність встановлюється для всієї синтаксеми, що характеризується єдністю і цілісністю значення, а не окремого слова, в яке закладено лише потенційні значення (у службових частин мови їх непомірно багато). І в цьому разі добре описана функція буде перекладачеві значно ліпшою підказкою, ніж намагання дати приблизний переклад.

Комп'ютерні розробки у галузі граматичної лексикографії ще більш раритетні, і серед них відзначається праця К. Маклеод, Р. Грішмана, А. Мейерса з укладання синтаксичного словника англійської мови для знарядь автоматичного опрацювання англійської мови [9]. У словнику передбачається детальний опис словосполучень, що можуть утворюватися з тими чи іншими частинами мови, деякі семантичні класи і розряди частин мови, компонентів словосполучення, хоча не описується власне семантика синтаксеми. Це й не дивно: для комп'ютерного опрацювання тексту не потрібне "людське" пояснення значення синтаксичної структури, тоді як користувачам мови воно кінцево необхідне. Окрім простого "людського пояснення", потрібні також описи синтаксем, і не лише у вигляді синтаксичних схем, а й у термінах граматичної і лексичної семантики; проблеми використання тремінологічного апарату граматичної семантики полягають у тому, що існує суттєвий різнобіч у номінації і наборах функцій і відношень у різних функціонально-граматичних школах.

У комп'ютерній лінгвістиці вказані напрями практично не розвиваються, незважаючи на вибух електронних технологій. Однак їх можна й доречно перетворити на неоціненні знаряддя комп'ютерної лексикографії. У комп'ютерній програмі пошук можна здійснювати як за реєстровим словом, так і за глосою, і за ілюстративним прикладом, як на початку зазначених позицій, так і всередині, і в кінці. Це надало б широкі можливості і всім тим, хто хоче повноцінно і якісно користуватися ресурсами мови, і професіоналам різних галузей: перекладачам, журналістам, юристам, широкому колу науковців. Отже, у сучасній комп'ютерній лексикографії формально-кількісний розвиток випереджає якісний. Тим не менш, ми переконані, що це явище тимчасове, і з часом виклики науки та вимога новизни спричиняться до розробки якісно нових словників. Словник синтаксем, який пропонуємо розглянути в статті, є саме одним з таких продуктів. Отже, мета статті – описати функції, параметри, особливості використання словника синтаксем та окреслити перспективи його оптимізації.

Матеріали і методи. Оскільки регулярна об'єктивна потреба використання такого словника виникла під час викладання дисципліни "комунікативні стратегії" для студентів освітньо-кваліфікаційного рівня "магістр", матеріалом для аналізу стали понад 200 синтаксем, що опрацьовуються у вправах під час практичних занять. Якщо перші роки спеціальної підготовки присвячені опануванню основних рівнів системи мови, на старших курсах саме час реорганізувати ці знання у функціонально-комунікативній парадигмі. Зазначені синтаксеми, зокрема, охоплюють такі поняттєві категорії: "аспектуальність" ("початковість", "однократність", "тривалість", "повторюваність", "завершеність", "результативність"), "атрибутивність", "безособовість", "гіпотетична модальність", "імперативність", "пасивність", "дебітивна модальність", "темпоральність", "умовність". Основні використані методи: семантичний і контекстуальний аналіз. Семантичний аналіз доречний тоді, коли аналізоване сполучення міститься у тлумачному словнику. Звичайно, не всі синтаксеми потрапляють до тлумачних словників, і тоді для виокремлення значення синтаксеми потрібно виходити з інтерпретації смислу висловлення у контексті. До прикладу, ствердна імперативність може виражатися такими синтаксичними моделями:

1) "*¡Anda!*" + *інфінітив*; 2) "*¡Venga!*" + *інфінітив*; 3) "*¡Hala!*" + *інфінітив*; 4) *Особовий займенник у функції непрямого додатка* + "*¡Importar!*" в особовій формі + *інфінітив*; 5) *дієслово "querer/preferir/necesitar" в особовій формі* + *сполучник "que"* + *підрядне додаткове речення з присудком у спосібі Modo Subjuntivo. та інші.*

Отже, словник може бути організований у вигляді ключових слів на позначення поняттєвих категорій, пов'язаних з ними семантичних нюансів, що супроводжуються посиланням до відповідної статті, мікροструктура якої для дескриптора "наказовість" міститься у попередньому абзаці.

Безсумнівно, потрібен і зворотний напрямок словника, а саме, від структури синтаксеми до її значення. До прикладу, синтаксеми, у складі яких міститься дієслово "acabar", синтаксичне оточення суттєво модифікує його значення. Проілюструємо прикладами оформлення мікροструктури:

ACABAR + DE + INFINITIVO 1) (*аспектуальність: завершеність*) *виражає дію, що щойно виконана; "щойно", "тільки що": "Los presidentes acaban de firmar la declaración";* 2) (*аспектуальність: результативність*) *вжитий з запереченням, виражає відсутність результату попри наявність дій, на нього спрямованих: "Por mucho que me expliques, no acabo de entender cómo puedes ser tan despistado".*

ACABAR + GERUNDIO (*аспектуальність: результативність*) *набути певної якості або стану з плином часу; "зрештою", "все-таки": "A fuerza de leer y copiar lo acabará comprendiendo".*

ACABAR + PARTICIPIO (*аспектуальність, пасивність*) *Опинитися у певному стані з плином часу внаслідок стороннього впливу або причини; "зрештою", "в результаті": "Después de la jornada acabó destrozado".*

ACABAR + EN + NOMBRE SUSTANTIVO (*аспектуальність: результативність*) *призвести до резуль-*

tamy “El desembarco del submarino acabó en tragedia”.

ACABAR +POR+ INFINITIVO (аспектуальність: результативність) домогтися результату за рахунок довготривалого впливу на щось; “зрештою”, “все-таки”: “Acabó por convencerme”.

NO ACABAR + DE + PARTICIPIO Див. *acabar + de + infinitivo*.

Отже, можна підсумувати, що потрібні принципово такі словники:

семантико-синтаксична структура – функції – відповідники функцій в іншій мові

Однак розробка словника такої структури може виявитися надто масштабним завданням на сьогодні, і для забезпечення усіх необхідних для перекладачів мовних пар знадобиться не одне десятиліття. Тому передбачене у схемі комплексне завдання можна цілком розподілити на дві незалежні ланки:

(I) синтаксична структура – опис функцій

(II) опису функцій – синтаксичні структури, що виконують описану функцію.

Перше з запропонованих завдань частково виконано як у загальних тлумачних словниках, так і в спеціалізованих, зазначених вище працях Е.В. Шевкун, В.І. Критської. Серед вітчизняних іспаністів, що працюють у галузі функціонального синтаксису в семасіологічному напрямку (від структури до змісту) виділяється праця “Синтаксис сучасної іспанської мови” Ю.А. Кочубинської-Смичковської [6].

Другий напрям ще не розроблено, і його можна охарактеризувати як створення ономаціологічного синтаксичного словника. Хоча можна без перебільшення стверджувати, що за на матеріалі української впритул наближені до цього праці І.Р. Вихованця. Зокрема, опис засобів вираження для функцій, що їх пропонує дослідник у “Нарисах з функціонального синтаксису української мови” [2] можна у широкому розумінні назвати ономаціологічним синтаксисом. У цій праці, як за задумом, так і за структурою, І.Р. Вихованець впритул наближується до розробки ономаціологічного синтаксичного словника. Безпосередньо у руслі ономаціологічного синтаксису працює В.П. Даниленко [5].

Результати та їх обговорення. Одна зі складнощів укладання словника синтаксем полягає у його неосязності. І.Р. Вихованець, К. Городенська при розробці граматичних словників дуже слушно зосередилися на службових частинах мови, оскільки в цих частиномовних категоріях з високим ступенем абстракції втілюються відношення між тими семантичними функціями, що виражаються повнозначними словами. У зв’язку з ідеєю граматичного словника для неслужбових частин мови К. Городенська зазначає: “Практики укладання граматичних словників на матеріалі повнозначних слів в українській мові ще немає. Причина зрозуміла: потрібен неабиякий хист, щоб у лексиког-

рафічному форматі викласти величезний обсяг граматичних функцій”. І не можна не погодитися, що, виходячи з критеріїв вичерпності, подібний словник був би неосязним, якщо тільки на прикладі іспанського дієслова “acabar” нам вдалося описати 6 синтаксем, і цей список ще не претендує на вичерпність.

У той самий час, якщо виходити з реальності і можливостей комп’ютерної галузі, розробку словника можна вести поступово, а критерієм виділення синтаксем цілком може стати їхня поширеність у репрезентативному корпусі малого обсягу. На вичерпність не претендує жодний словник: так само, як жоден лексикограф не подумає включити до словника всі до останнього неологізми, як та зенонівська черепаха, яку ніколи не наздожене Ахіллес (поки друкуватиметься словник, з’являться нові слова), можна стверджувати, що всеосязно врахувати загальний зріз синтаксем – теж завдання непідйомне. Приміром, у словнику *DRAE (Diccionario de la Real Academia Española)*, найдетальнішому сучасному тлумачному словнику іспанської мови, у статті заголовним словом “acabar” відсутні пояснення щодо керування дієслова прийменником “en + іменник” зі значенням “приводити до результату”, у той час як корпус письмове та усне мовлення рясніє прикладами вживання цієї синтаксеми.

Висновки. Підсумовуючи, зазначимо, що словник синтаксем потрібен широкому колу фахівців: практикам (перекладачам, викладачам, журналістам), і теоретикам, позаяк він надає можливість віднаходити потрібні засоби вираження для тих чи інших змістів, відношень, а також урізноманітнювати їх у мовленні. Поза тим, у такому словнику можна аналізувати зв’язки і відношення, в яких перебувають між собою синтаксеми, поняттєві категорії, як однієї, так і різних мов. Отже, словник може становити інтерес сам собою як об’єкт лінгвістичного конструювання: його характеристики, властивості, формальні й кількісні параметри можуть і мають стати предметом досліджень.

Двомовні і багатомовні словники синтаксем видаються перспективними, однак, з огляду на складність їх укладання, більш реалістичним видається укладання двомовних словників за двома можливими схемами:

1) від опису синтаксеми у термінах граматичних категорій та функцій граматичної семантики до набору значень синтаксеми (семасіологічний словник);

2) від можливого значення до набору синтаксем, що можуть його виражати (ономаціологічний словник).

Параметризація словника за допомогою знарядь комп’ютерної лексикографії може суттєво розширити можливості пошуку, а також дозволити об’єднувати словники синтаксем двох або більше мов (семасіологічного й ономаціологічного) в одному двомовному або багатомовному словнику.

ЛІТЕРАТУРА

1. Вихованець І.Р. Теоретична морфологія української мови / І.Р. Вихованець, К.Г. Городенська. – Київ: Наукова думка, 1980. – 288 с.
2. Вихованець І.Р. Нариси з функціонального синтаксису української мови / І.Р. Вихованець. – Київ: Наукова думка, 1992. – 224 с.
3. Городенська К.Г. Граматичний словник української мови: сполучники / К.Г. Городенська. – Київ-Херсон: Видавництво ХДУ, 2007. – 340 с.
4. Городенська К. Граматична лексикографія: стан і перспек-

- тиви / К.Г. Городенська // Українська мова : науково-теоретичний журнал. – Київ, 2018. – № 1 (65). – С. 30-37.
5. Даниленко В.П. Ономазиологическое направление в грамматике. 2-е изд / В.П. Даниленко. – Москва, 2007. – 344 с.
6. Кочубинська-Смичковська Ю.А. Синтаксис сучасної іспанської мови [Текст] : навч. посібник / Ю. Кочубинська-Смичковська. - Київ: ВПЦ "Київський університет", 2000. - 175 с.
7. Критська В.І. Граматичний словник української літературної мови : словозміна : близько 140 000 слів / Критська

- В.І. [та ін.]. – Київ: Вид. дім Дмитра Бураго, 2011. – 757 с.
8. Шевкун Е.В. Іспансько-російсько-український словник граматичних лексем. – Київ: Освіта України, 2014. – 284 с.
9. Macleod. C., Grishman. R., Meyers. A. Creating a common syntactic dictionary of English // Proceedings of the International Workshop on Sharable Natural Language Resources, 1994. – [Електронний документ]. – Режим доступу від 22.11.2018: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.44.1933>

REFERENCES

1. Vykhoanets I.R. Theoretical morphology of the Ukrainian language / I.R. Vykhoanets, K.H. Horodenska. – Kyiv: Naukova dumka, 1980. – 288 p.
2. Vykhoanets I.R. Studies on functional syntax of the Ukrainian language / I.R. Vykhoanets. – Kyiv: Naukova dumka, 1992. – 224 p.
3. Horodenska K.H. Grammatical dictionary of the Ukrainian language: conjunctions / K.H. Horodenska. – Kyiv-Kherson: State University of Kherson Ed., 2007. – 340 c.
4. Horodenska K.H. Grammatical lexicography: state and perspectives / K.H. Horodenska // The Ukrainian Language: a theoretical science journal. – Kyiv, 2018. – № 1 (65). – P. 30-37.
5. Danilenko V.P. Onomasiological branch in grammar. II Edition / V.P. Danilenko. – Moscow, 2007. – 344 p.
6. Kochybynska-Smychkovska Y.A. Modern Spanish Syntax / Y. A. Kochybynska-Smychkovska. - Kyiv: Kyiv Taras Shevchenko University Ed., 2002. – 175 p.
7. Krytska V.I. Grammatical dictionary of the Ukrainian standard language: wordforms: around 140.000 lexemes / Krytska V.I. and others ed. by N.F. Klymenko. – Kyiv: Ed. Office “Dmytro Burago”, 2011. – 757 p.
8. Shevkun E.V. Spanish-Russian-Ukrainian dictionary of grammatical lexemes. – Kyiv: Osvita Ukrainy, 2014. – 284 p.
9. Macleod. C., Grishman. R., Meyers. A. Creating a common syntactic dictionary of English // Proceedings of the International Workshop on Sharable Natural Language Resources, 1994. Retrieved November 22, 2018 from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.44.1933>

Electronic dictionary of syntaxemes: microstructure’s organization and possible applications

S. B. Fokin

Abstract. Nowadays lexicography is characterized by essentially new products, and among them, grammar oriented dictionaries. Though grammar dictionary may be compiled according to different patterns and for different purposes, those, particularly useful for a wide range of practical and scientific-theoretical usages are syntaxeme dictionaries, in which syntactic structures along with their meaning should be represented, both in semasiological and onomasiological ways. Once compiled, a clever structuring of dictionary’s database is to facilitate all kind of formal search throughout all the parts of its macro- and microstructure, as well as merging different languages dictionaries into a bilingual one.

Keywords: *syntaxeme, lexicography, onomasiology, electronic dictionary, grammatical semantics, functional syntax, translation.*