

О. І. ПОПОВА, Є. В. КУПРІЯНОВ

ЕЛЕКТРОННИЙ СЛОВНИК АНГЛІЙСЬКОЇ ДІЄСЛІВНОЇ ДІЄВІДМІНИ: ЙОГО БУДОВА, ЗМІСТ І ПАРАМЕТРИЗАЦІЯ

Стаття присвячена розв'язанню головних питань створення електронного словника англійської дієслівної дієвідміни з використанням теорії лексикографічних систем (Л-систем). Зокрема, визначено формальну структуру англійського дієслова, схарактеризовано її квазіоснови та квазіфлексії, а також побудовано модель Л-системи словника та визначено його лексикографічні структури, до яких передбачається доступ для користувача. Визначено лексикографічні параметри для репрезентації словозмінної парадигми дієслова. Окреслено завдання навчального та дослідницького характеру, які може розв'язувати розроблюваний словник.

Ключові слова: Л-система, англійська мова, дієвідміна, комп'ютерна лексикографія, словозміна, електронний словозмінний словник.

О. И. ПОПОВА, Е. В. КУПРИЯНОВ

ЭЛЕКТРОННЫЙ СЛОВАРЬ АНГЛИЙСКОГО ГЛАГОЛЬНОГО СЛОВОИЗМЕНЕНИЯ: ЕГО СТРУКТУРА, СОДЕРЖАНИЕ И ПАРАМЕТРИЗАЦИЯ

Статья посвящена решению главных вопросов создания электронного словаря английского глагольного спряжения, используя теорию лексикографических систем (Л-систем). В частности, определена формальная структура английского глагола, охарактеризованы квазиосновы и квазифлексии, а также построена модель Л-системы словаря и определены его лексикографические структуры, которые предполагается сделать доступными для пользователя. Оговорены лексикографические параметры для описания глагольной парадигмы. Определены задачи учебного и исследовательского характера, которым может решать разрабатываемый словарь.

Ключевые слова: Л-система, английский язык, спряжение, компьютерная лексикография, словоизменение, электронный словоизменятельный словарь.

О. І. ПОПОВА, YE. V. KUPRIYANOV

ELECTRONIC DICTIONARY OF ENGLISH VERB CONJUGATION: ITS STRUCTURE, CONTENTS AND PARAMETRIZATION

The article studies the main issues related to creating electronic dictionary of English verb conjugation by using the theory of lexicographic systems (L-systems). In particular the formal structure of the English verb was determined together with its possible quasi-stems and quasi-flexions. To facilitate functioning of the inflexion dictionary in digital environment the L-system model was built up and lexicographic structures to be made accessible for user were determined. Lexicographic parameters to represent verb conjugation were specified. The tasks of educational and research nature which are possible to be resolved with the English verb conjugation dictionary in question were set out.

Keywords: L-system, English, conjugation, computational lexicography, inflexion, electronic inflexion dictionaries.

Вступ. Розроблюваний електронний словник англійської дієслівної дієвідміни передбачається використовувати не лише як інформаційно-довідковий ресурс для користувачів, а й також для систем опрацювання текстової інформації. У нашій роботі мова іде про користувацький варіант словника, призначений для виконання навчальних функцій. Зауважимо, що в Інтернеті наявна чимала кількість таких ресурсів, що подають парадигму для англійського дієслова, зокрема: *ReversoConjugation*, *bab.la*, *Verbix*, *Vocabulix* та інші. Але, на жаль, їх навчальний функціонал доволі обмежений, тобто лише показують парадигму для обраного дієслова. Але, на наш погляд, цього недостатньо, оскільки користувача може цікавити інформація іншого плану, наприклад: які ще дієслова, крім аналізованого, мають такий самий принцип словозміни; чи наявні групи дієслів, що можуть зазнавати морфологічних змін під час дієвідмінювання; чи можуть слова утворювати варіативну парадигму тощо.

Дієслово в граматичній побудові англійської мови займає абсолютно виняткове місце внаслідок широкої розгалуженості своєї системи і тієї великої ролі, яку дієслівні форми грають в побудові пропозиції. Як знаменна частина мови має категоріальне значення «процесуальності»: дієслово передає процес, що динамічно розвивається в часі, включаючи не тільки власне дії (*to work*, *to build*), але також стану, форми

існування (*to be*, *to become*, *to lie*), різні види відносин, почуття (*to love*, *to appreciate*) тощо.

У граматиці [10] англійські дієслова класифікують, головним чином, за їхніми морфологічними, лексико-морфологічними, синтаксичними та функціональними властивостями. Деякі лінгвісти пропонують лексичну класифікацію [11–15]. Пропоновані класифікації безумовно є цінними. Для нас вони є основою для розроблення принципів формального моделювання системи англійської дієслівної словозміни.

Аналіз останніх досліджень і публікацій. Питання формального моделювання системи словозміни та її представлення у програмних засобах (віртуальні лексикографічні лабораторії, електронні граматичні словники) детально висвітлено в працях директора Українського мовно-інформаційного фонду НАН України, академіка В. А Широкова [4–7] та його співпрацівників Т. П. Любченко [2] та І. В. Шевченка [3]. Варто зазначити, що першим, хто почав працювати над проблемами моделювання системи словозміни (іменної) та її лексикографічного упорядкування, є А. А. Залізник. Його наукова праця «Русское именное словоизменение» і «Грамматический словарь русского языка» лягли в основу теорії Л-систем граматичного типу. Запропонована теорія системи мови В. А. Широкова успішно реалізовано під час моделюванні системи словозміни української [3], російської [3], німецької [2] та турецької мов [8].

Зокрема, досліджено: 1) їх мовний інвентар (флексії, квазіфлексії) для утворення парадигми; 2) критерії класифікації мовних одиниць за їхніми словозмінними властивостями; 3) характер утворення парадигми (повна / неповна); 4) параметри словозміни тощо. Але отримані напрацювання для зазначених мов неможливо застосувати автоматично у моделюванні системи англійської словозміни.

Мета дослідження. Виходячи зі сказаного, мета нашої розвідки – розглянути окремі теоретичні питання побудови електронного словника англійської дієслівної словозміни. Для цього необхідно: 1) подати формальну модель англійського дієслова, виокремити його квазіоснову та квазіфлексію; 2) побудувати модель Л-системи словника англійської дієслівної дієвідміни; та 3) схарактеризувати лексикографічні параметри для репрезентації дієслівної дієвідміни. Теоретичне підґрунтя складають праці В. А. Широкова та його співпрацівників [4–7].

Матеріали і результати дослідження. Моделювання англійської дієслівної словозміни починається зі встановлення та формалізації лінгвістичних критеріїв, за якими вся множина дієслів розбивається на певні підмножини, взаємний перетин яких є порожнім і всередині кожної з них словозміна відбувається за одним правилом. Підмножини дієслів з такими властивостями називаються граматичними групами.

У межах граматичних груп формуються граматичні класи, що представляють певні різновиди словозміни. Так, в англійській мові ми виокремлюємо дві граматичні групи:

- *Правильні дієслова*, що дієвідмінюються за стандартними правилами;
- *Неправильні дієслова*, у яких відбуваються певні відхилення від правил або утворення індивідуальної словозміни.

У свою чергу, другу граматичну групу можна розкласти на граматичні класи, що представляють словозміну з певними правилами, наприклад, подвоєння або чергування кінцевої приголосної тощо. Наприклад, граматична група «Неправильні дієслова» поділяється на такі класи:

- *Дієслова з індивідуальною словозміною*, тобто дієслова із суплетивними формами;
- *Дієслова з незмінюваною формою* у формах минулого часу дійсного способу;
- *Дієслова з 2 або 3 однаковими формами* у формах Simple Past і Past Participle.

Останній граматичний клас містить також варіативну парадигму в формі минулого часу (Simple Past). Особливості поділу граматичних типів на граматичні групи типу показано на рисунку 1.

Щоб забезпечити можливість функціонування словника в цифровому середовищі, потрібно розробити його формальну модель, що відображає структурні елементи словника, а також відношення між ними. Під її розробляння концептуальної моделі, ми врахували типи лінгвістичної інформації, яку має надавати розроблюваний словник: 1) словозмінну

парадигму для будь-якого дієслова; 2) морфологічну характеристику будь-якої словоформи; 3) вихідну форму будь-якого слова, взятого з тексту.

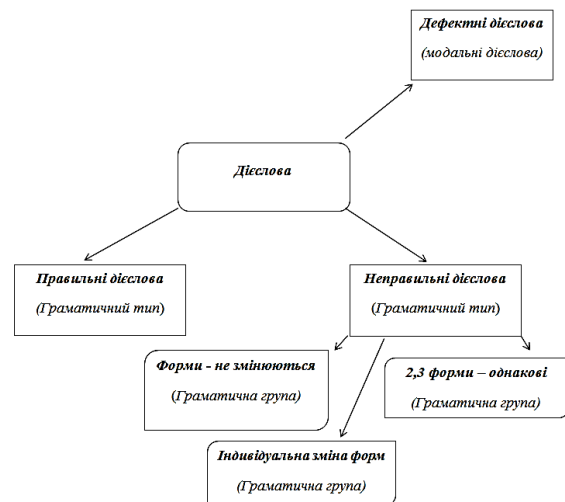


Рис. 1. – Парадигматичні класи та групи дієслова

Отже, структуру словника можна представити в такому формалізованому вигляді (у вигляді Л-системи):

$$\{I^D(PAR), V(PAR), \beta, \sigma[\beta], H^{PAR}, H^{LEM}\}, \quad (1)$$

де $I^D(PAR)$ – сукупність усіх мовних одиниць іспанської мови, $V(PAR)$ – множина лексикографічних описів, β – елементи лексикографічного опису, $\sigma[\beta]$ – структури, утворені на $V(PAR)$, H^{PAR} та H^{LEM} – оператори парадигматизації та лематизації. За методологією теорії Л-систем, граматичну лексикографічну систему можна представити, як це показано на рис. 1.

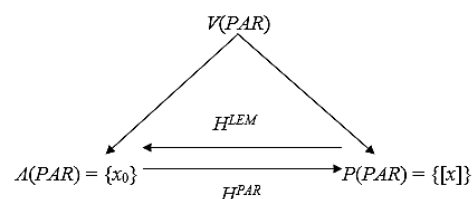


Рис. 2. – Загальна структура Л-системи

У наведеній схемі вказані елементи структури Л-системи мають таку інтерпретацію: 1) $V(PAR)$ – множина словникових статей; 2) $A(PAR)$ – множина лівих частин словникових статей, що містить одиниці в канонічній формі (інфінітиву) $\{x_0\}$; 3) $P(PAR)$ – множина правих частин словникових статей, яку складають всі словозмінні парадигми $\{[x]\}$; 4) H^{PAR} – оператор парадигматизації, що дає змогу побудувати парадигми для будь-якого слова, заданого канонічною його формою x_0 ; 5) H^{LEM} – оператор лематизації, що перетворює будь-який елемент парадигми $[x]$ у канонічну форму x_0 . Розглянемо детально структуру x та $[x]$.

За теорією Л-систем граматичного типу, слово можна розкласти на постійну і змінювану частину. Першу називають квазіосною, а другу – квазіфлексією. Формально таке визначення ми можемо представити таким чином:

$$x = c(x) * f(x), \quad (2)$$

де x – іспанська мовна одиниця в певній граматичній формі, а $c(x)$ та $f(x)$ – відповідно, квазіоснова та квазіфлексія. Квазіфлексії можуть охоплювати такі частини слова:

- $c(x) = 0$, якщо слово має суплетивну форму, як наприклад, англійське дієслово *to be* (бути) у другій та третій особі однини теперішнього часу дійсного способу: *am, is, are*;
- $f(x) = 0$, якщо компонент x є незмінюваним. Порівняйте, наприклад, дієслово *to put* в усіх особових формах неозначеного минулого часу, пасивного дієприкметника;
- $c(x) = x - f(x)$, якщо x має змінювану квазіфлексію, як, наприклад, у дієслова *to steal* в особових формах минулого часу (*steal – stole*) та пасивного дієприкметника (*steal – stolen*).

У загальному значенні словозміною певної мови називають відповідність, за якою кожному слову відповідає його парадигма [3, с. 3]. Отже, перетворивши рівняння (1) відповідно до визначення А. А. Залізняка, отримуємо формальний вигляд словозміни:

$$[x] = c(x) * [f(x)], \quad (3)$$

де $[x]$ – сукупність слів форм (парадигма) одиниці x , а $[f(x)]$ – набір квазіфлексій, що входять до складу парадигми $[x]$. Символом $*$ позначено конкатенацію квазіоснови з квазіфлексією. Вважають, що під час словозміни тотожність слова (лексеми) не порушується (тобто ми маємо справу з тим самим словом в різних граматичних формах) на відміну від словотвору, де від одного слова можуть утворюватися інші, відмінні від нього слова. Кожна дієслівна парадигма $[x]$ описується набором інформаційних елементів, що складають множини $V(PAR)$.

Уважаємо необхідним уточнити, що множини $V(PAR)$ складають елементи, які в теорії Л-систем називають β -структурами, а набір значень, що можуть вони набувати – доменами. Перелічимо домени β -структур:

- 1) $Dom\beta_1 = \{x_0 \in X\}$ – дієслова у вихідній формі;
- 2) $Dom\beta_2 = \{P_1, P_2, \dots, P_n\}$ – домени граматичних груп;
- 3) $Dom\beta_3 = \{G^1, G^2, \dots, G^{N(P_i)}\}$ – домени граматичних класів, утворюваних усередині кожній граматичній групі;
- 4) $Dom\beta_4 = \{\Pi^1, \Pi^2, \dots, \Pi^{N(G_i)}\}$ – домени парадигматичних класів, належних певним граматичним класам;
- 5) $Dom\beta_5 = \{\omega^i_0, \omega^i_1, \dots, \omega^i_{N(P_i)}\}$ – домени словозмінних параметрів англійських дієслів;

6) $Dom\beta_6 = \{f^k_0, f^k_1, \dots, f^k_{N(G_i)}\}$ – набори квазіфлексій відповідних парадигматичних класів Π .

Структура $\lambda[\beta]$ лівої частини словникової статті визначається множиною кортежів, заданих на декартовому добутку $Dom\beta_1 \times Dom\beta_2 \times Dom\beta_3 \times Dom\beta_4$. Таким чином, структура лівої частини має формат <слово у вихідній формі>, <граматичний клас>, <граматична група>. Структура $\rho[\beta]$ правої частини визначається множиною кортежів, заданих на декартовому добутку $Dom\beta_4 \times Dom\beta_5 \times Dom\beta_6$. Формат структури правої частини є такий: <парадигматичний клас>, <граматичне значення>, <квазіфлексія>.

Виходячи з вищесказаного, словникова стаття (див. рис. 3) матиме таку структуру: 1) заголовкове слово, тобто дієслово у вихідній формі; 2) таблиця парадигм, що подає усі особові та безособові форми; та 3) словозмінний класифікатор, що вказує на граматичну групу і граматичний клас, до яких належить обране дієслово.

[заголовкове слово] – $Dom\beta_1$	
[таблиця парадигм] – $Dom\beta_4 \times Dom\beta_5 \times Dom\beta_6$	
Словозмінний класифікатор – $Dom\beta_2 \times Dom\beta_3 \times Dom\beta_4$	
Граматична група	[назва]
Граматичний клас	[назва]
Парадигматичний клас	

Рис. 3. – Схема словникової статті

Словник словозміни дієслів англійської мови подає морфологічну інформацію заданого дієслова, зокрема:

1) всі форми теперішнього (Present Simple, Progressive, Perfect, Perfect – Progressive), минулого (Past Simple, Progressive, Perfect, Perfect – Progressive) та майбутнього (Future Simple, Progressive, Perfect, Perfect – Progressive) часів, для двох чисел (однини та множини) та для трьох осіб (перша – I / you, друга – you / you, третя – he / she / it / they);

2) форми наказового (Imperative mood) та умовного (Subjunctive mood) способу;

3) безособові форми: інфінітив (Infinitive), дієприкметник (Participle) та герундій (Gerund).

Додаткова функція словника – пошук дієслова за класом та групою. Так, наприклад, якщо користувач обирає клас «Правильні дієслова», то результатом пошуку стане перелік дієслів, розподілених по групах, які формуються залежно від характеру морфологічних змін під час додавання *-ed*. Наприклад, якщо основа дієслова закінчується: а) на приголосну і *-y*, то остання змінюється на *-i* (*cry – cried – cried*); б) на *-e*, то після додавання *-ed* пишеться тільки одна буква *-e* (*translate – translated – translated*), в) на ударну коротку голосну і приголосну після неї, то у 2 і 3 формі згодна в кінці слова подвоюється (*admit – admitted – admitted*).

Висновки. Розроблений словник англійської дієслівної дієвідміни передбачається використовувати під час лінгвістичних досліджень та навчального процесу. Щодо перших, є можливість досліджувати: 1) закономірність змін, що відбуваються під час утворення парадигми; 2) вплив лексичного значення на реалізацію іспанськими одиницями словозмінної парадигми; 3) явища омографії, тобто одиниць, що збігаються одна з одною за своєю граматичною формою, але належать до різних парадигматичних класів. У навчальних цілях словник може: а) подавати парадигму для конкретного слова, б) надавати перелік слів, що відмінюються за однаковими правилами в) вказувати на відмінності словозміни слів, що належать до одного граматичного класу або групи. Розроблені моделі й структура словника уможливають використання словника у програмах для автоматичного формування повної словозмінної парадигми для будь-якої реєстрової одиниці словника (формування всіх можливих варіантів граматичних форм), у програмах автоматичного морфологічного аналізу текстових словоформ і лематизації довільної текстової словоформи.

Список літератури

1. Зализняк А. А. *Русское именное словоизме.*: Наука, 1967. 372 с.
2. Любченко Т. П. *Лексикографічні системи граматичного типу та їх застосування в засобах автоматизованого опрацювання мови*: автореф. дис. ... канд. техн. наук: 10.02.21. К, 2011. 19 с.
3. Любченко Т. П., Шевченко І. В., Широков К. В. Граматичні лексикографічні системи з флективним і аглютинативним компонентом словозміни. *Мовознавство*. 2013. № 2–3. С. 209–220.
4. Широков В. А. Граматика у феноменологічному вимірі. *Мовознавство*. 2014. № 4. С. 3–27.
5. Широков В. А. *Інформаційна теорія лексикографічних систем*. К.: Довіра, 1998. 331 с.
6. Широков В. А. *Комп'ютерна лексикографія*. К.: Наукова думка, 2011. 351 с.
7. Широков В. А. *Феноменологія лексикографічних систем*. К.: Наукова думка, 2004. 328 с.
8. Широков К. В. *Іменна словозміна у сучасній турецькій мові*. К.: Довіра, 2009. 318 с.
9. *Языкознание. Большой энциклопедический словарь*. Гл. ред. В. Н. Ярцева. 2-е изд. М.: Большая Российская энциклопедия, 1998. 685 с.
10. Eastwood J. *Oxford Guide to English Grammar*. Oxford: Oxford University Press, 2002. 453 p.
11. Kipper, K., A. Korhonen, N. Ryant, and M. Palmer. Extending VerbNet with Novel Verb Classes. In: *Proceedings of the 5th International Conference on Language Resources and Evaluation*. Genova, Italy. 2006a. pp. 1027–1032.
12. Kipper-Schuler, K. *VerbNet: A broad-coverage, comprehensive verb lexicon*. Ph.D. thesis, Computer and Information Science Dept., University of Pennsylvania, Philadelphia, PA. 2005. 158 p.
13. Korhonen, A. and N. Ryant. *Novel Lexical-Semantic Verb Classes*. Available at: verbs.colorado.edu/~kipper/Papers/euralex-final.doc (accessed 20.03.18)
14. Korhonen, A. and Briscoe. T. Extended Lexical-Semantic Classification of English Verbs. In: *Proceedings of the HLT/NAACL Workshop on Computational Lexical Semantics*. Boston, MA. 2004.
15. Levin B. *English Verb Classes and Alternation, A Preliminary Investigation*. The University of Chicago Press, 1993. 366 p.

References (transliterated)

1. Zaliznjak A. A. Russkoe imennoe slovoizmenenie [Russian nominal inflection]. Moscow, Nauka, 1967. 372 p.
2. Lyubchenko T. P. *Leksykohrafichni systemy hramatychnoho typu ta yih zastosuvannya v zasobah avtomatyzovanoho opracyuvannya movy [Lexicographical systems of grammar type and their application in the automated language processing tools]*. Kyiv, 2011. 19 p.
3. Ljubchenko T. P., Shevchenko I. V., Shyrokov K. V. Gramatychni leksykohrafichni systemy z flektyvnyvm i agljutynatyvnyvm komponentom slovozmyny [Grammar lexicographical systems with flexional and agglutinative component of inflection]. In: *Movoznavstvo*. 2013. no. 2-3, pp. 209–220.
4. Shyrokov V. A. Gramatyka u fenomenologichnomu vymiri [Grammar in phenomenological dimension]. In: *Movoznavstvo*. 2014. no. 4. pp. 3–27.
5. Shyrokov V. A. *Informacijna teorija leksykohrafichnyh system [Information theory of lexicographical systems]*. Kyiv, Dovira, 1998. 331 p.
6. Shyrokov V. A. *Komp'uterna leksykohrafija [Computer lexicography]*. Kyiv, Naukova dumka, 2011. 351 p.
7. Shyrokov V. A. *Fenomenologija leksykohrafichnyh system [Phenomenology of lexicographical systems]*. Kyiv, Naukova dumka, 2004. 328 p.
8. Shyrokov K. V. *Imenna slovozmyna u suchasnij turec'kij movi [Nominal inflection in the modern Turkish language]*. Kyiv, Dovira, 2009. 318 p.
9. *Jazykoznanie. Bol'shoj jenciklopedicheskij slovar' [Linguistics. Comprehensive Encyclopedic dictionary]*. Ed. V. N. Yartseva. 2nd edition. Moscow, Bol'shaja Rossijskaja jenciklopedija, 1998. 685 p.
10. Eastwood J. *Oxford Guide to English Grammar*. Oxford: Oxford University Press, 2002. 453 p.
11. Kipper, K., A. Korhonen, N. Ryant, and M. Palmer. Extending VerbNet with Novel Verb Classes. In: *Proceedings of the 5th International Conference on Language Resources and Evaluation*. Geneva, Italy. 2006. pp. 1027–1032.
12. Kipper-Schuler, K. *VerbNet: A broad-coverage, comprehensive verb lexicon*. Ph.D. thesis, Computer and Information Science Dept., University of Pennsylvania, Philadelphia, PA. 2005. 158 p.
13. Korhonen, A. and N. Ryant. *Novel Lexical-Semantic Verb Classes*. Available at: verbs.colorado.edu/~kipper/Papers/euralex-final.doc (accessed 20.03.18)
14. Korhonen, A. and Briscoe. T. Extended Lexical-Semantic Classification of English Verbs. In: *Proceedings of the HLT/NAACL Workshop on Computational Lexical Semantics*. Boston, MA. 2004.
15. Levin B. *English Verb Classes and Alternation, A Preliminary Investigation*. The University of Chicago Press, 1993. 366 p.

Надійшла (received) 27.03.2018

Відомості про авторів / Сведения об авторах / About the Authors

Попова Оксана Ігорівна (Попова Оксана Игоревна, Popova Oksana Ihorivna) – Національний технічний університет «Харківський політехнічний інститут», студентка факультету соціально-гуманітарних технологій; м. Харків, Україна; e-mail: oksannna5@gmail.com

Курпянов Євген Валерійович (Курпянов Евгений Валерьевич, Kurpyanov Yevhen Valeriyovych) – кандидат філологічних наук, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри інтелектуальних комп'ютерних систем; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-0801-1789>; e-mail: eugeniokuprianov@gmail.com