

The methodological foundation of the research is made up of the ideas propagated by O.O. Selivanova, G. Cook, P. Stockwell and E. Semino.

The main result of the research is demonstrating the interaction between purely narratological phenomena and the methods of cognitive poetic analysis.

**Key words:** cognitive map of a literary discourse, module of the addresser, module of the text, module of the addressee, conceptual space, blending, possible worlds, fallible narrator.

Стаття надійшла до редколегії  
14.03.2013 р.

УДК 811.111'322+811.111'42

Ірина Біскуб

### Кванти лінгвістичної інформації в лексичних базах даних

У статті здійснено спробу ввести в науковий обіг термін «квант лінгвістичної інформації». Такий підхід дає змогу впритул підійти до встановлення принципів формального структурування різних типів для комп'ютерної їх обробки. Запропоновано аналіз квантів лінгвістичної інформації в лексичній базі даних DANTE.

**Ключові слова:** когнітивна категоризація, категорія, лінгвістична категорія, лінгвальні знання.

**Постановка наукової проблеми та її значення.** Однією з найактуальніших проблем сучасної когнітивної лінгвістики є вироблення чітких і науково верифікованих методик дослідження вербальної діяльності мозку людини за допомогою формального символічного опису неформальних структур. 1995 р. Р. Джекендофф увів поняття «ментальних репрезентацій» [3], за допомогою яких він пропонує експліцитно виражати ідеальні структури мозку. Протягом останніх декад учені намагаються розробити моделі інтелектуальної діяльності людини, спираючись на метафоричне, але цілком логічне розуміння мозку як системи обробки інформації, так званого «біологічного комп'ютера» [3, с. 1].

Дослідників найбільше цікавить, які саме типи інформації здатен обробити мозок людини, тобто які категорії, властивості, зв'язки може опрацювати мозок, і як із їх допомогою формально описати поведінку, життєвий досвід та інтелект.

**Мета статті** – спроба ввести в науковий обіг термін «квант лінгвістичної інформації», за допомогою якого формально структуруватиметься лінгвістична інформація у комп'ютерних лексичних базах даних.

**Виклад основного матеріалу й обґрунтування отриманих результатів дослідження.** На сучасному етапі відкриваються широкі перспективи щодо інтегративного застосування лінгвістичних категорій для формального опису лінгвальних та екстралінгвальних знань. Прихильники когнітивного розуміння категоризації стверджують, що когнітивні категорії безпосередньо пов'язані з категоріями лінгвістичними, а «засвоєння лінгвістичних категорій впливає на процес утворення концептуальних нелінгвістичних категорій» [6, с. 256]. Лінгвістичні категорії – це «лінгвістичні одиниці, що вживаються у лексичній, морфологічній та синтаксичній системах мови» [5, с. 403]. О. С. Ахманова визначає лінгвістичні категорії (англ. *linguistic categories*) як «загальні властивості різноманітних класів і розрядів мовних одиниць, що отримують різноманітне мовне вираження (морфологічне, синтагматичне, просодичне та ін.) із різним змістом (граматичним, лексичним, стилістичним та ін.)» [1, с. 599].

На думку вчених, до складу лінгвістичних категорій входять елементи, що поводяться як група, демонструючи спільні дистрибутивні зразки, спільні способи вживання та спільні обмеження. Так, традиційно граматичні категорії представлені класами слів (іменник, дієслово, прикметник та ін.). Морфологічні категорії, наприклад, специфікують маркери узгодження (*agreement markers*) [5, с. 403]. До складу синтаксичних категорій входять типи іменникових, дієслівних конструкцій, типи речень (*noun phrases, verb phrases, types of sentences*) [5, с. 403]. Ці категорії демонструють спільні ознаки в

багатьох мовах. Незважаючи на наявні внутрішньо системні відмінності, існує багато подібностей у «лінгвістичній поведінці» іменникових словосполучень при їх протиставленні дієслівним словосполученням (Noun Phrase vs Verb Phrase).

Однак, як зазначає Дж. Лакофф, і мовна категоризація, і лінгвістична категоризація є когнітивно орієнтованими процедурами, в основу яких покладено принципи когнітивного моделювання знань, як-от:

- 1) пропозиціоналізація – специфікація елементів, їх властивостей і зв'язків згідно з структурною організацією знань у мозку людини;
- 2) створення образних схем – схематична специфікація властивостей (наприклад, довжина, товщина, контейнери й ін.);
- 3) метафоризація – специфікація пропозиціональних і схематичних властивостей одних елементів за допомогою подібних властивостей інших елементів;
- 4) метонімізація – представлення одного елемента як моделі іншого елемента [4].

Усі проаналізовані підходи до розуміння когнітивної категоризації передбачають розробку стандартів до формального опису лінгвістичної інформації, оскільки саме за умови врахування особливостей ментального структурування знань звичайний словник як основний репозитарій лінгвістичних знань трансформується у базу даних із можливістю комп'ютерної обробки.

Розглянемо детальніше принципи формального кодування лінгвістичної інформації у лексичній базі даних DANTE. Лексична база даних DANTE була створена 2011 р. на основі корпусу сучасної англійської мови обсягом 1,7 млн слововживань. Структура бази даних була розроблена Б. С. Аткинсом із урахуванням її теорії лексикографічної релевантності [2, с. 252], що базується на ретельному формалізованому аналізі конкордансів понад 50 тис. слів сучасної англійської мови.

Розглянемо моделі структурування лінгвістичної інформації у згаданій базі даних. На рисунку 1 наведено видобуток із бази даних щодо ключового слова *computer*.

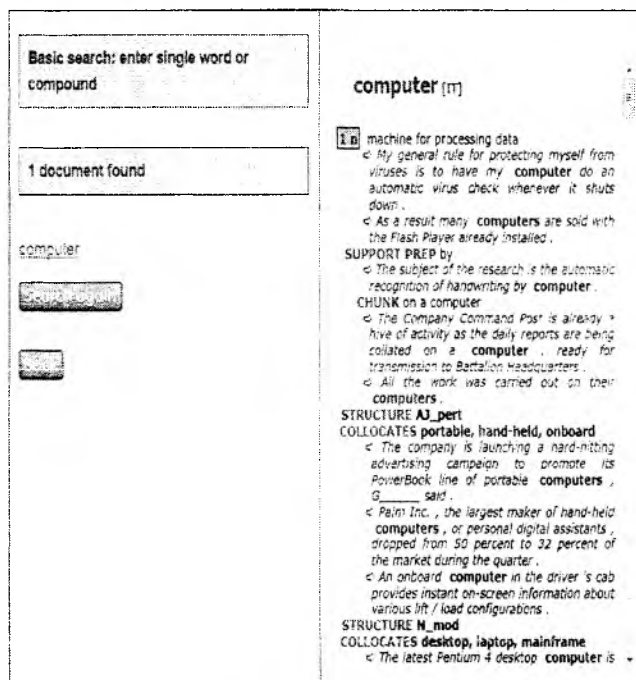


Рис. 1. Видобуток із DANTE для лексеми *computer*

Для кожної лексики DANTE робить видобуток із електронного корпусу за двома основними принципами:

- 1) видобуток лексичних одиниць, що асоціюються з лемою;
- 2) додавання релевантних квантів лінгвістичної інформації до кожної видобутої лексичної одиниці.

Асоційовані лексичні одиниці видобуваються через виокремлення контекстуальних вжитків (senses) лематизованих слів або словосполучень. У структурі бази даних вжитки пронумеровані. Нумерація базується на частотності вживання, тобто найчастотніші вжитки представлені на початку списку. Наприклад, для лєми *computer* маємо таку ієрархію вживань у мовленні:

- 1) *n. machine for processing data;*
- 3) *n. computer technology in general;*
- 4) *n. SUBFORM computers: data-processing machines as an area of interest or expertise;*
- 5) *adj. done, achieved with the help of computers.*

Заслугує на увагу той факт, що ранжування вживань лексичних одиниць не є статичною інформацією. У цьому полягає основна відмінність видобутку із бази даних порівняно із видобутком з електронного словника. Щоразу, задаючи пошук у базі даних, отримуємо ранжування контекстуальних значень лексичної одиниці відповідно до реального стану речей у сучасному мовленні. Тобто, якщо слово втрачає частотність в одному з варіантів свого вживання, це автоматично відобразиться і на його порядковому номері у результатах видобутку без жодного втручання лексикографа.

Що стосується формального розпізнавання релевантних квантів лінгвістичної інформації, то для цього проаналізуємо перелік лінгвістичних категорій, які експліцитно представлені у лексичній базі даних DANTE. Під *квантом лінгвістичної інформації* розуміємо формально виражену й комп'ютерно марковану парцелу лінгвістичної категорії, яка є маркером поля для комп'ютерної бази даних. У таблиці 1 наведено 12 квантів лінгвістичної інформації, за якими маркуються елементи бази даних.

Таблиця 1

Кванти лінгвістичної інформації як маркери полів бази даних DANTE

Field	Example of attributes
part of speech: 16 attributes	noun, verb, adjective, adverb etc.
inherent grammatical properties: 32 attributes	<ul style="list-style-type: none"> <li>• nouns: the count / uncount distinction, the unitary / mass distinction, pluralization etc.</li> <li>• verbs: impersonal, passive-only etc.</li> <li>• adjectives: attributive, predicative, post-modifying etc.</li> <li>• adverbs: manner, degree, viewpoint, and sentence adverbs</li> </ul>
valency / syntactic contexts: nouns: 18 attributes	modified by or modifying a noun; clausal, infinitival and prepositional complements etc.
valency / syntactic contexts: verbs: 42 attributes	intransitive, transitive and ditransitive uses; clausal, infinitival and prepositional complements etc.
valency / syntactic contexts: adjectives: 17 attributes	premodified by noun or adverb; clausal, infinitival and prepositional complements etc.
linguistic labels: domain:	accountancy, aerospace, agriculture, airforce, anatomy etc.

Field	Example of attributes
156 attributes	
linguistic labels: region: 6 attributes	American English, British English, Australian English etc.
linguistic labels: register: 4 attributes	formal, informal, very informal, vulgar
linguistic labels: style: 21 attributes	child language, drug abuse language, euphemism, journalese etc.
linguistic labels: time : 2 attributes	dated, obsolete
linguistic labels: attitude : 3 attributes	appreciative, offensive, pejorative.

Як засвідчує аналіз даних таблиці, комп'ютер здатен розпізнавати 11 парцел лінгвістичної інформації із їх подальшим дефрагментуванням на категоріальні кванти, серед яких:

- 1) part of speech;
- 2) inherent grammatical properties;
- 3) valency / syntactic contexts: nouns;
- 4) valency / syntactic contexts: verbs;
- 5) valency / syntactic contexts: adjectives;
- 6) linguistic labels: domain;
- 7) linguistic labels: region;
- 8) linguistic labels: register;
- 9) linguistic labels: style;
- 10) linguistic labels: time;
- 11) linguistic labels: attitude.

Найчисленнішою є парцела «*domain*», яка репрезентована 156 варіантами квантів лінгвістичної інформації на позначення галузей знань, до яких належить вжиток тієї чи іншої лексичної одиниці. Привертає також увагу факт синтаксичного тлумачення лінгвістичної категорії валентності, яка на синтагматичному рівні мовленнєвого вжитку традиційно репрезентується категорією сполучуваності. Цілком закономірно, що валентнісна парадигма є найпоширенішою для дієслова, оскільки англійська мова вербоцентрична, що засвідчено і кількістю дієслівних категоріальних квантів.

**Висновки.** Наведений аналіз категоріального квантування лінгвальних знань у електронній лексичній базі даних, з одного боку, засвідчує порівняно обмежену кількість категоріальних квантів, які здатен розпізнавати комп'ютер. З іншого боку, сама можливість такого розпізнавання підтверджує релевантність сучасної лінгвістичної теорії з огляду на потенціал комп'ютерного кодування лінгвістичних категорій та їх автоматичну обробку. Саме у розширенні кількості комп'ютерно упізнаваних категоріальних квантів сучасної англійської мови вбачаємо перспективу подальших досліджень у галузі когнітивної категоризації.

#### Джерела та література

1. Ахманова О. С. Словарь лингвистических терминов / О. С. Ахманова. – М. : Сов. энцикл., 1969. – 607 с.
2. Atkins B. T. S., Fillmore Ch., Johnson Chr. (2003). Lexicographic relevance: selecting information from corpus evidence / B. T. S. Atkins, Ch. Fillmore, Chr. Johnson // International Journal of Lexicography, guest editor Thierry Fontenelle. Oxford, OUP: 16:3, 2003. – P. 251–280.

3. Jackendoff R. Languages of the Mind: Essays on Mental Representations / R. Jackendoff. – Cambridge, MA : A Bradford Book, the MIT Press, 1995. – 200 p.
4. Lakoff G. Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought / G. Lakoff, M. Johnson. – Chicago : University of Chicago Press, 1999. – 624 p.
5. Lillo-Martin D. Syntactic categories in Signed Versus Spoken Language / D. Lillo-Martin // Handbook of Categorization in Cognitive Science // H. Cohen, C. Lefebvre (eds.). – Elsevier LTD, 2005. – P. 401–421.
6. Papafragou A. Realations between Language and Thought: Individuation and the Count / Mass Distinction / A. Papafragou // Handbook of Categorization in Cognitive Science // H. Cohen, C. Lefebvre (Eds.). – Elsevier LTD, 2005. – P. 255–275.

**Бискуб Ирина.** Кванты лингвистической информации в лексических базах данных. В статье изложены научно-практические аргументы для введения термина «квант лингвистической информации». Предложенное новое понятие кванта лингвистической информации дает возможность вплотную подойти к проблеме формального структурирования лингвистической информации с целью ее компьютерного кодирования в электронных лексических базах данных. Квант лингвистической информации определяем как формально выраженную и компьютерно маркированную парцеллу лингвистической информации, которая является маркером поля в компьютерной базе данных. На примере маркирования полей лексической базы данных современного английского языка DANTE проанализированы основные кванты лингвистической информации, которые может автоматически распознавать компьютер.

**Ключевые слова:** когнитивная категоризация, категория, лингвистическая категория, лингвальные знания.

**Biskub Iryna.** The Quanta of Linguistic Information in Lexical Databases. The article presents scientific and practical reasons for introducing the new term «quantum of linguistic information». The notion of the quantum of linguistic information gives way to approaching the problem of discriminating formal structures for decoding linguistic information. Such decoding enables computational processing of linguistic information in lexical databases. The quantum of linguistic information is formal and computationally explicit fragment of linguistic information, which serves as a field marker in lexical database. The field markers of the DANTE Lexical Database of Modern English have been analyzed as the basic quanta of linguistic information to be recognized and processed by computer.

**Key words:** cognitive categorization, category, linguistic category, linguistic knowledge.

Стаття надійшла до редколегії  
19.03.2013 р.

УДК 811.112.2-373

**Марія Блажко**

### **Ад'єктивно-іменні словосполучення на позначення статичних джерел звуку: метонімічний ракурс (на матеріалі сучасної німецької мови)**

У статті концептуалізація відношень між статичним джерелом звуку та його властивостями пояснюється за допомогою метонімії як внутрішньоконцептної процедури репрезентації звучання. Встановлено, що звучання осмислюється за допомогою просторової й об'єктної метонімії. Просторова метонімія сприяє осмисленню ситуації продукування звуку джерелом, яке перебуває в певному місці, а об'єктна метонімія відбиває розпізнавання джерела звуку як цілого за його ознаками або частинами, втілюючись у трьох моделях сполучуваності прикметників та дієприкметників на позначення звучання з іменниками на позначення частин тіла й артефактів.

**Ключові слова:** ад'єктивно-іменні словосполучення, ад'єктивні одиниці на позначення звучання, концептуальна метонімія, просторова метонімія, об'єктна метонімія.

**Постановка наукової проблеми та її значення.** Лінгвістичні розвідки останніх років спрямовані на дослідження мовних засобів вербалізації знань та уявлень про світ і мову, набутих внаслідок взаємодії людини з навколишнім середовищем. Застосування семантико-когнітивного підходу до вивчення семантики та комбінаторики прикметників і дієприкметників на позначення звучання в