

О.В. Бояринцева

ЛІНГВОСТАТИСТИЧНІ ЗАЛЕЖНОСТІ МОВОЗНАВСТВА (НА МАТЕРІАЛІ АРАБОМОВНОЇ КОМП'ЮТЕРНОЇ ТЕРМІНОЛОГІЇ)

Систему знань певної науки становлять виявлені факти, їхній поняттєвий, якісний і кількісний опис, а також емпіричні закономірності, що були встановлені шляхом їхнього аналізу. Для цілісного наукового уявлення про дійсність необхідно визначити те загальне або спільне, яке стосується всього світоутворення чи окремих його частин, тобто закон або групу законів. Отже, відкриття та формулювання законів становить чи не найголовнішу мету наукових досліджень. З другого боку, проаналізувавши місце законів у загальній системі теоретичного знання, можна зробити висновок, що закон являє собою серцевину, ядро окремої наукової теорії, а система законів – основу певної науки в цілому. У сучасній науці закон виступає важливим складовим елементом наукової теорії [Кочерган 2006]. Наявність специфічних для певної науки законів свідчить про досягнення цією наукою високого рівня розвитку.

Таким чином, і мовознавство, як наука дуже давня і розгалужена (як стверджується у більшості підручників та посібників з мовознавчих наук), мало б вже досягти певного високого рівня розвитку, одним із показників якого, як ми визначили раніше, є наявність законів. Та, якщо ми запитасмо пересічного студента філологічної спеціальності про мовознавчі або лінгвістичні закони, навряд чи отримаємо вичерпну відповідь. Що, у свою чергу, вказує на серйозні недоліки сучасної української філологічної освіти, які треба терміново виправляти. Бо ж мовознавчі закони існують – вони відкриті, доведені, навіть класифіковані. Незрозумілим є нехтування цими законами. Отже, проблема лінгвістичних законів сьогодні не лише не втра-

чає своєї **актуальності**, а й постає особливо гостро.

Поняття „лінгвістичного закону” можна визначити таким чином: деяке загальне правило або загальна закономірність, характерна для певної мови, різних мов або мови взагалі; регулярне й послідовне відтворення того чи іншого співвідношення одиниць даної мови, що можна уявити у вигляді формули закономірних співвідношень. Становлення лінгвістики як самостійної науки нерозривно пов'язане з пошуком закономірностей в існуванні й розвитку конкретних природних мов, людської мови взагалі та наміром відобразити кожен закономірність у вигляді лінгвістичного закону.

Сьогодні відкрито й досліджено майже дві сотні лінгвістичних закономірностей та законів (різноманітні розподіли, детермінативні послідовності, кореляції, константи). Безцінним у цій сфері лінгвістичних знань є доробок професора К.М. Тищенка. Вже відкриті мовні закономірності, класифіковані ним за п'ятьма типами (розподіли Зіпфа, розподіли Гауса, детермінативні послідовності, кореляції та константи) і розподілені за шістьма метатеоретичними галузями мовознавства (загальнолінгвістичні закони, закономірності контрастивно-порівняльної лінгвістики, закономірності аксіологічної лінгвістики, закономірності ідіолоктології, закономірності синтактики й закономірності сигніфіки) [Тищенко 2007].

Загальною метою цього дослідження є спроба привернути увагу до несправедливо проігнорованих більшістю українських мовознавців лінгвістичних законів та закономірностей і переконатися в їхній дії на матеріалі арабомовної комп'ютерної

та інтернет-термінології (яка є предметом майбутнього дисертаційного дослідження). Виконання **поставленого завдання** передбачає використання статистичного методу дослідження та аналіз отриманих у результаті дослідження даних.

Дія лінгвістичних законів і закономірностей в арабській мові раніше розглядалася лише в бакалаврській роботі на тему „Частотний словник сучасної арабської мови” студента Інституту філології Київського національного університету ім. Т.Г. Шевченка Б. Бабака. Проте у певних термінологічних системах арабської мови такі дослідження раніше не проводилися. **Перспективи подальших** досліджень у зазначеному напрямку для арабської мови надзвичайно широкі з огляду на її певні семантичні та морфологічні особливості, про що буде йти мова нижче.

Закономірності, що діють у терміносистемах, належать до закономірностей аксіологічної лінгвістики, тобто закономірних особливостей будови, функціонування й історичних змін етномовних аксіолектів: діалектів, говірок, вернакул – у процесі мовленнєвої комунікації. До аксіолектів належать і терміносистеми різноманітних галузей людських знань. Серед основних закономірностей аксіопрагматики можемо назвати такі:

- розподіл Зіпфа в жанрах, функціональних стилях і терміносистемах (розподіл лексики в текстах відрізняється залежно від жанру або стилю; розподіл термінологічної лексики в наукових текстах відповідає залежності Зіпфа);
- розподіл Лотки (розподіл вчених за числом публікацій, цитувань чи згадувань у підручниках);
- закономірність Бредфорда (ранговий розподіл журналів з кількістю статей на певну тематику);
- правило Асколі (там, де найгустіша сітка діалектів, там прабатьківщина мови);
- правило периферійних архаїзмів (явище, що існує в кількох острівних ареалах на периферії території певної етномови, в минулому було поширене на всій її території).

Розглянемо деякі лінгвостатистичні закономірності (як аксіопрагматики, так і синтактики, про які мова піде пізніше) на конкретних прикладах. Об'єктом дослідження буде арабомовна комп'ютерна та інтернет-термінологія. Підрахунки проводилися, зокрема, на матеріалі статті „Kaufa nata'āmalu ma'a talfi al-aqrāṣ aṣ-ṣulba wa-wiqāyatuha” (“Виправлення пошкоджень жорстких дисків та запобігання їм”), опублікованої в інтернет-версії щоденного незалежного кувейтського видання „Аль-Кабас” (дата публікації у друкованому виданні – 23 березня 2008 р.). Загальна кількість слів – 2201. Розподіл перших найчастотніших слів – *див. табл. 1* (Двадцять найчастотніших слів досліджуваної статті).

Виділивши з усього масиву вживаної у статті лексики лише термінологічну, маємо такі дані – *див. табл. 2* (П'ятнадцять найчастотніших термінів досліджуваної статті).

І в першій, і в другій таблицях очевидно є залежність рангу слова від частоти його вживань, що є ілюстрацією основного закону Есту – Зіпфа. Основний закон Есту – Зіпфа формулюється так: „Добуток частоти слова на його ранг є величина стала”, тобто частота слова обернено пропорційна його рангу – номеру в частотному вокабулярі до тексту. Основний закон Есту – Зіпфа ще називається першим правилом Зіпфа [Тищенко 2007].

Серед розподілів Зіпфа також можна назвати такі закономірності синтактики:

- друге правило Зіпфа (чим слово частотніше, тим воно коротше);
- третє правило Зіпфа (чим слово частотніше, тим воно багатозначніше);
- залежність Зіпфа – Гердана (чим слово частотніше, тим воно сполучуваніше);
- четверте правило Зіпфа (чим слово частотніше, тим воно давніше);
- закономірність Бумера в паузах вагання (у спонтанному мовленні найбільше пауз між реченнями, і вони є найдовшими, потім між групою підмета і групою присудка);
- закон Зіпфа – Гіро (чим слово менш частотне, тим таких слів більше);
- залежності Зіпфа у словотворі (при

рівномірно прибувному n кількість n -членних словотвірних гнізд у словнику спадає).

Графік рангового розподілення слів за частотою їхнього вживання – *див. графік 1*. Цей графік ілюструє перше правило Зіпфа. Оскільки масштаб осі ординат логарифмічний, ранговий характер розподілу відображає спадаюча крива лінія. Кут нахилу кривої для різних масивів лексики може бути різним залежно від особливостей стилю автора.

Погляньмо на перші десять найчастіше вжитих слів у статті: чотири з них складаються з однієї літери (wa-, bi-, fa-, li-), ще чотири – з двох (min, fi, huwa, hiya), і два – з трьох літер (qāḡṣ, ṣulb), що прекрасно ілюструє друге правило Зіпфа (чим слово частотніше, тим воно коротше).

Розглянемо тепер значення найбільш вживаних слів. Так, прийменник bi- має вісім основних груп значень, сполучник wa – шість, прийменник min – шість, прийменники fi, fa- та li- – по сім, чотири та п'ять відповідно, іменник qāḡṣ – шість, прикметник ṣulb – два основні значення, його високий ранг зумовлений темою статті, а займенники huwa та hiya хоч і мають по одному семантичному значенню, проте зайняли високий ранг, оскільки у підрахунках були враховані всі їхні похідні форми, форми непрямих відмінків. Отже, очевидна залежність частотності слова від його багатозначності (третє правило Зіпфа).

Доволі часто частота вживання різних слів збігається. Так, перші чотирнадцять рангів посідають окремі слова, а вже п'ятнадцятий та шістнадцятий ранги посили по три слова кожен. Починаючи ж із 25-го рангу, всі наступні посідають чотири і більше слів. При цьому кількість слів, що зайняли один ранг, зростає (*див. табл. 3* (Залежність рангу слова і кількості слів цього рангу)). Найнижчий ранг займає найбільша кількість слів. Про що й говорить закон Зіпфа – Гіро (чим слово менш частотне, тим таких слів більше).

У процесі написання статті стало зрозуміло, що дослідження дії лінгвістичних законів та закономірностей на матеріалі

арабської мови має ряд своїх особливостей. Перш за все мова йде про слова, що пишуться разом зі словами, до яких вони відносяться. Так, сполучники wa- та fa- пишуться разом з наступним словом, прийменники bi- та li- пишуться разом зі словом, до якого відносяться. Цілий ряд (12!) особових займенників пишуться разом з іншими словами залежно від значення, що ними передається. Чи розділяти такі „сполуки”? І якщо розділяти, то як саме? Виразно постає проблема слова та словоформи. Як у цьому контексті розглядати форми ламаної множини в арабській мові? Форми непрямих відмінків займенників? Чи враховувати якимсь чином у частотності вживань артикль? У вирішенні всіх вищенаведених питань полягають перспективи подальшого дослідження дії лінгвостатистичних законів в арабській мові.

Під час проведення підрахунків, наведених у цій статті, ми розглядали як одне слово різні форми ламаної множини та однини, різні форми непрямих відмінків, проте виділяли як окремі слова прийменники, займенники та сполучники, що пишуться разом зі словами, до яких відносяться; не враховували артикль.

Для розв'язання вищезазначених проблем необхідно насамперед визначити поняття слова в арабській мові. Застосування інших критеріїв роботи зі словами, звичайно, певним чином змінить отримані нами результати, проте загальної картини дії лінгвістичних законів та закономірностей не змінить.

Після проведення дослідження можемо зробити такі **висновки**: лінгвостатистичні залежності мовознавства можуть бути проілюстровані на матеріалі арабомовної комп'ютерної термінології; дія лінгвостатистичних законів для арабомовної комп'ютерної термінології, порівняно з лексикою загального вжитку, має свої певні особливості, що зумовлені як семантичними та морфологічними характеристиками арабської мови, так і специфікою термінологічного тексту; подальші дослідження у зазначеному напрямку вимагають вирішення низки теоретичних питань.

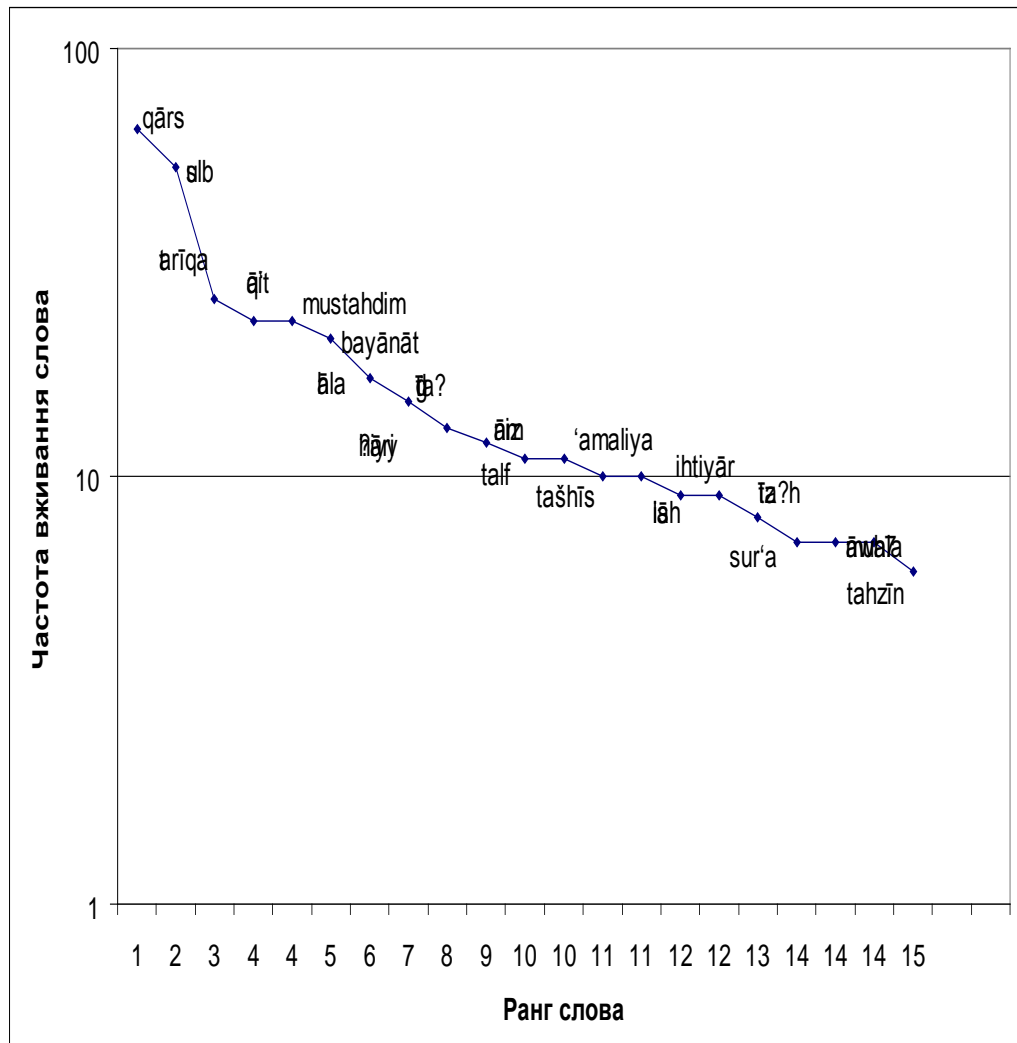
Таблиця 1. Двадцять найчастотніших слів досліджуваної статті

Ранг слова	Слово	Значення слова	Частота вживання	Відсоток покриття тексту
1	bi-	у, з, при	129	5,86%
2	wa-	і	128	15,35%
3	min	з	81	
4	huwa	він	69	21,44%
5	qārş	диск	65	
6	fī	в	63	32,84%
7	hiya	вона	57	
8	şulb	жорсткий	53	
9	li-	для	43	
10	fa-	та	35	
11	aw	чи	33	
12	‘ala	на	32	
13	tilka	та, той	28	
14	ṭarīqa	спосіб	26	
15	qāma	виконувати	25	
16	qīṭā‘ mustahdim yumkin	сектор, кластер користувач могти	23	48,79%
17	bayānāt tah‘iyya ḥāsūb	дані підготовка комп'ютер	21	
18	allati naw‘	яка, який вид	20	
19	inna	що (сполучник)	18	
20	ḥāla	стан, спосіб	17	

Таблиця 2. П'ятнадцять найчастотніших термінів досліджуваної статті

Ранг слова	Слово	Значення слова	Частота вживання	Відсоток покриття тексту
1	qārş	диск	65	2,95%
2	şulb	жорсткий	53	6,54%
3	ṭarīqa	спосіб, шлях, призначення	26	
4	qīṭā‘ mustahdim	сектор, кластер користувач	23	9,58%
5	bayānāt	дані	21	
6	ḥāla	спосіб, стан	17	13,17%
7	taşgīl	запуск, робота	15	
8	ḥāriḡiyyu	зовнішній	13	
9	niẓām	система	12	
10	talf ‘amaliya	пошкодження процес	11	
11	taşhīş	розпізнавання	10	
12	iḥtiyār işlāh	вибір виправлення	9	16,49%
13	taḡhīz	підготовка	8	
14	sur‘a muḥāwala mu‘āliḡ	швидкість спроба процесор	7	
15	tahzīn	резервне збереження	6	

Графік 1. Рангове розподілення слів за частотою їхнього вживання



Таблиця 3. Залежність рангу слова і кількості слів цього рангу

Ранг слова	Частота вживань	Кількість слів, що посіли ранг
1–15	129, 128, 81, 69, 65, 63, 57, 53, 43, 35, 33, 32, 28, 26 і 25 відповідно	1
16–17	23 і 21 відповідно	3
18	20	2
19–20	18 і 17 відповідно	1
21	16	2
22	15	4
23–24	14 і 12 відповідно	1
25	11	5
26	10	4
27	9	6
28	8	5
29	7	11
30	6	12
31	5	10
32	4	12
33	3	36
34	2	77
35	1	228

ЛІТЕРАТУРА

- Баранов Х.К. **Арабско-русский словарь**. Москва, 2001.
- Кочерган М.П. **Загальне мовознавство**. Київ, 2006.
- Тищенко К.М. **Закони мовознавства: чи потрібні вони науці?** (http://www.langs.com.ua/movy/zakony/zakony_mov.htm)
- Тищенко К.М. **Метатеорія мовознавства**. Київ, 2000.
- Тищенко К.М. **Основи мовознавства**. Київ, 2007.
- Філіпченко А.С. **Основи наукових досліджень**. Київ, 2005.
- Хайтун С.Д. **Наукометрия. Состояние и перспективы**. Москва, 1983.