

СЛІПЕЦЬКА В. Д.

*Дрогобицький державний педагогічний університет
імені Івана Франка*

ЛЕКСИЧНА ДОВЖИНА АНГЛІЙСЬКИХ ІНТРАГАЛУЗЕВИХ ТЕРМІНІВ ШТУЧНОГО ІНТЕЛЕКТУ, МАТЕРІАЛОЗНАВСТВА, АЕРОКОСМІЧНИХ ТЕХНОЛОГІЙ ТА ФІНАНСОВО-ЕКОНОМІЧНОЇ СПРАВИ

У статті йдеться про лексичну, графічну довжину інтрагалузевих термінів галузей штучного інтелекту, матеріалознавства, аерокосмічних технологій та фінансово-економічної справи. В усіх чотирьох англійських термінологіях розподіл термінів за лексичною, графічною довжиною має багато спільного: 50 % термінів припадає на двокомпонентні одиниці, 10 % – на однокомпонентні одиниці. З-поміж чотирьох термінологій дещо вирізняється термінологія аерокосмічних технологій. Це єдина термінологія у якій ми зафіксували семи- та восьмикомпонентні одиниці. Взаємозв'язок між лексичною та графічною довжиною термінів та частотою їх використання у текстах підлягає закону Ципфа: коротші терміни вживаються частіше.

Ключові слова: інтрагалузевий термін, взаємозв'язок, лексична довжина, графічна довжина, частота терміна.

В статье исследуется лексическая, графическая длина интраотраслевых терминов искусственного интеллекта, материаловедения, аэрокосмических технологий и финансово-экономической отрасли. В четырех терминологиях распределение терминов за лексической, графической длиной имеет много общего. Двухкомпонентные термины составляют 50 %, однокомпонентные – 10 %. Из четырех терминологий терминология аэрокосмических технологий несколько отличается. Это единственная терминология в которой функционируют семи-, восьмикомпонентные единицы. Взаимосвязь между лексической и графической длиной интраотраслевых терминов и их частотой функционирования в текстах объясняется законом Ципфа: короткие термины используются чаще.

Ключевые слова: интраотраслевой термин, взаимосвязь, лексическая длина, графическая длина, частота термина.

The article focuses on the lexical, graphical lengths of intradomain terms that function in English terminologies of artificial intelligence, metal processing, aerospace technologies, finance and economics. The division of intradomain terms according to lexical, graphical lengths has much in common: two-component terms constitute 50 %, one-component terms constitute 10 %. The English terminology of aerospace technologies differs from three other terminologies. Seven-component, eight-component terms function in texts that belong to this terminology. There exists the correlation between the term length and its frequency. This correlation is explained by Zipf's Law: short terms are of high frequency.

Key words: intradomain term, correlation, lexical length, graphical length, term frequency.

У вітчизняній та зарубіжній лінгвістиці термінознавчі дослідження можна поділити на дві групи: роботи загальнотеоретичного характеру, в яких розглядається специфіка терміна, його властивості тощо [1], та описи термінологій окремих

галузей [2; 3]. Проте комплексне зіставлення різних галузевих термінологій в межах однієї мови не виконувалося. **Актуальність** дослідження визначається загальною тенденцією сучасного мовознавства, спрямованою на зіставні студії. Йдеться про зіставлення різних термінологічних систем в межах тієї самої мови.

Науковий текст з будь-якої тематики містить інтрагалузеві терміни “спеціальні терміни тієї галузі знання, якій належить текст” [4:15]. Не є винятком і тексти зі штучного інтелекту, матеріалознавства, аерокосмічних технологій, фінансово-економічної справи. Однією з вимог до терміна є лаконічність [5; 6]. Це спонукає нас до врахування такої ознаки, як довжина терміна. Під лексичною довжиною терміна ми розуміємо кількість кореневих морфем повнозначних слів. Довжина лексичної одиниці, зокрема слова, у контексті синергетичної лінгвістики визнається важливим чинником самоорганізації мовної системи [7]. Вона корелює з частотою слова, його віком, коефіцієнтом полісемії. Досліджується зв'язок довжини слова з такими когнітивними характеристиками, як швидкість розпізнавання його смислу [8], запам'ятовування та пригадування [9] тощо. У стилеметричних працях активно досліджується зв'язок довжини слова із жанром та / чи авторством текстів [10; 11]. Довжина терміна корелює з його семантичною складністю та інформаційною насиченістю [12]. Семантична складність терміна тут розглядається в аспекті моносемії. Вважається, що чим більший інтенціонал значення, тим семантично складніший термін (і більша його інформаційна насиченість), пор. *image : color image : computer generated image (1)*.

Мета дослідження – з'ясувати спільні та відмінні риси довжини інтрагалузевих термінів штучного інтелекту, матеріалознавства, аерокосмічних технологій, фінансово-економічної справи.

Об'єкт – інтрагалузеві терміни зі штучного інтелекту, матеріалознавства, аерокосмічних технологій, фінансово-економічної справи.

Предмет – лексична, графічна довжина інтрагалузевих термінів.

Емпіричну базу дослідження складають терміни, отримані шляхом суцільної вибірки з наукових статей (теоретичного характеру) технічного та гуманітарного напрямів. Загальний обсяг опрацьованого матеріалу – 5851 термін (за галузями, відповідно: штучний інтелект – 1499, матеріалознавство – 1466, аерокосмічні технорлогії – 1436, фінансово-економічна справа – 1450).

Інтенціонал значення слова чи (терміна) можна експлікувати шляхом аналізу відповідної дефініції. Проте це досить трудомістка операція, а у нашому випадку практично нездійсненна, оскільки більшість термінів штучного інтелекту, які функціонують у сучасних наукових текстах, не зафіксована у словниках (принаймні, доступних нам). Тому саме через структурну складність термінів, зокрема їхню довжину, матимемо змогу хоч би приблизно оцінювати семантичну складність та інформаційну насиченість термінів.

Так само, як довжину слова можна вимірювати по-різному – кількістю фонем, графем, складів, морфем [13], по-різному можна визначати і довжину терміна. Найчастіше враховують кількість складів та літер. Е.Ф.Скороходько запропонував розгорнуту номенклатуру оцінок довжини терміна: синтаксичну (кількість слів у складі терміна), лексичну (кількість кореневих морфем), словотвірну (загальну кількість морфем – кореневих і словотвірних) та графічну (кількість літер) [14].

Лексична довжина інтрагалузевих термінів у чотирьох досліджених термінологіях складає в середньому 2,381 компонента, за галузями: штучний інтелект – 2,275, аерокосмічні технології – 2,728, матеріалознавство – 2,248, фінансово-економічна справа – 2,262. У трьох випадках (штучний інтелект, матеріалознавство, фінансово-економічна справа) різниця несуттєва (статистично незначуща). Проте аерокосмічні терміни демонструють більшу, порівняно з трьома іншими галузями, довжину (ймовірність помилки $p=0,000000$). Таким чином, можна стверджувати, що, принаймні для деяких галузей, довжина може слугувати дистинктивною ознакою термінологій.

Крім середньої лексичної довжини, було обчислено також розподіл термінів за цим параметром, тобто кількість одиниць різної довжини. Табл. 1 презентує розподіл термінів чотирьох досліджуваних термінологій (штучний інтелект *AI*, аерокосмічні технології *AS*, матеріалознавство та обробка металу *MP*, фінансово-економічна справа *FE*) за довжиною, тобто за кількістю кореневих морфем *L*. Позначення у таблиці: *N* – кількість термінів відповідної довжини у вибірці, *Sum* – накопичена кількість, тобто сумарна кількість термінів, які мають таку чи меншу довжину, % – частка термінів у відсотках до загальної кількості інтрагалузевих термінів відповідної дисципліни, *C%* – накопичена частка термінів.

Таблиця 1

Розподіл термінів за довжиною у різних галузях

L	AI				AS			
	N	Cum	%	C%	N	Cum	%	C%
1	111	111	9,9	9,9	80	80	7,0	7,0
2	665	776	59,5	69,4	475	555	41,7	48,8
3	279	1055	25,0	94,4	383	938	33,7	82,4
4	53	1108	4,7	99,1	120	1058	10,5	93,0
5	7	1115	0,6	99,7	50	1108	4,4	97,4
6	3	1118	0,3	100	16	1124	1,4	98,8
7					12	1136	1,1	99,8
8					2	1138	0,2	100

L	MP				FE			
	N	Cum	%	C%	N	Cum	%	C%
1	150	150	12,8	12,8	139	139	13,8	13,8
2	683	833	58,4	71,2	549	688	54,5	68,3
3	255	1088	21,8	93,0	246	934	24,4	92,8
4	62	1150	5,3	98,3	63	997	6,3	99,0
5	20	1170	1,7	100	9	1006	0,9	99,9
6	1	1174	0,1	100	1	1007	0,1	100

В усіх чотирьох термінологіях розподіл термінів має багато спільного. Близько половини термінів припадає на двокомпонентні одиниці, чверть – на трикомпонентні, приблизно десять відсотків – на однокомпонентні терміни. Схожість розподілу особливо наочно виявляється на графічному зображенні (рис. 1).

З-поміж чотирьох термінологій дещо вирізняється за розподілом аерокосмічна. В ній відносно менше одно- та двокомпонентних одиниць, натомість більше довших термінів. Крім того, це єдина термінологія, у якій ми зареєстрували семи- та восьмикомпонентні терміни, наприклад: *aerolastic tiltrotor wind-tunnel model*, *airplane-mode aerolastic stability testing*, *augmented frequency-dependent pressure-coupled combustion response*, *four-bladed semi-articulated soft-inplane rotor system* (2).

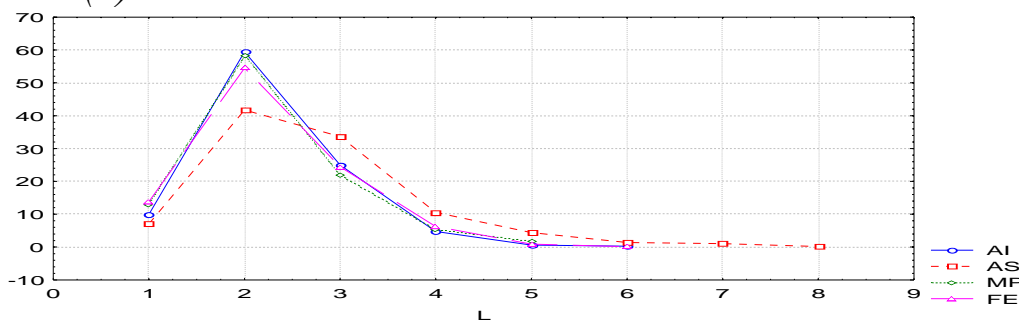


Рис. 1. Розподіл спеціальних термінів за довжиною

Однак слід констатувати, що тенденції розподілу залишаються незмінними. Криві розподілу решти термінологій майже збігаються. Таким чином, враховуючи, що чотири досліджені терміносистеми належать до різних за характером галузей знання, можна зробити висновок, що розподіл термінів за лексичною довжиною – це одна з термінологічних універсалій.

Схожий вигляд має також розподіл за графічною довжиною – кількістю літер. На рис. 2 відображено узагальнений розподіл термінів усіх чотирьох термінологій за цим параметром.

Розподіл термінів як і за лексичною довжиною (кількістю кореневих морфем), так і за графічною довжиною (кількістю літер) суттєво відрізняється від розподілу мовних одиниць згідно з відомим законом Ципфа. За цим законом чим коротша одиниця, тим більша її частота. Щоправда, це стосується, в першу чергу, частоти у тексті, а вище йшлося про частоту термінів у системі мови, у словнику. Якщо врахувати частоту термінів у текстах, характер їх розподілу за довжиною дещо змінюється – частота однокомпонентних термінів збільшується, а багатокомпонентних – зменшується. Проте і в цьому разі найчастотнішими залишаються двокомпонентні терміни.

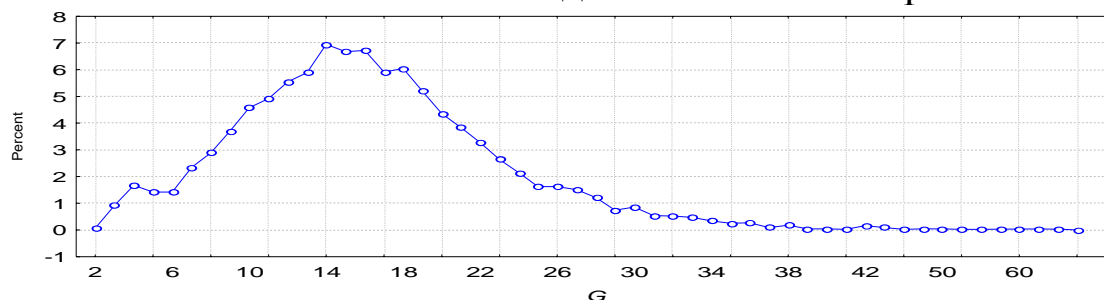


Рис. 2. Розподіл термінів за графічною довжиною

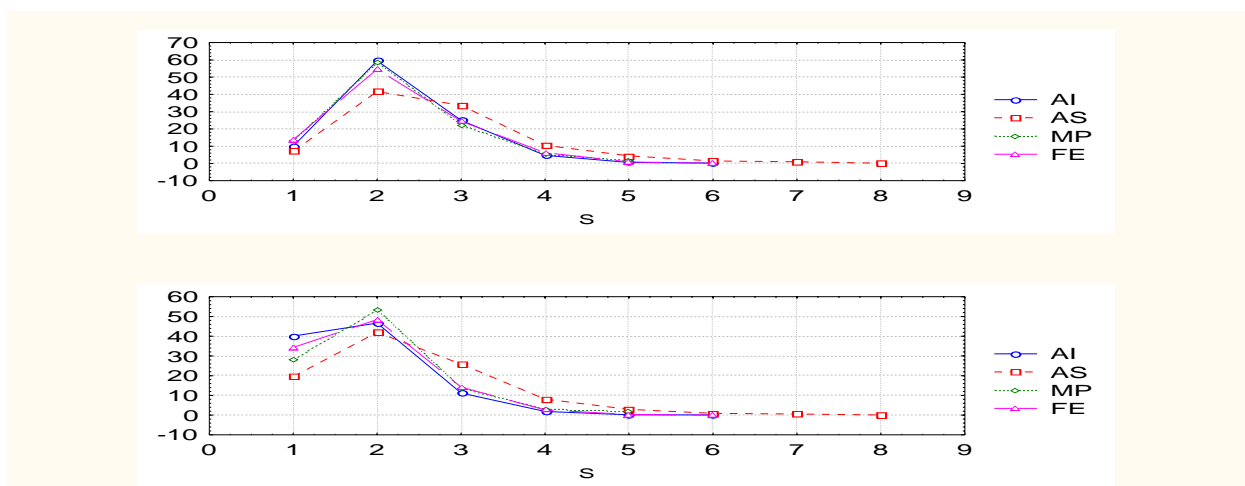


Рис. 3. Розподіл термінів за довжиною у словнику і у тексті

На рис. 3 зображено розподіл термінів за довжиною у словнику (верхні графіки) та у тексті (нижні графіки). Ймовірна причина відхилення наших даних від закону Ципфа полягає ось у чому. Ципф у своїх спостереженнях виходив з принципу найменших зусиль (у Мартіне – принцип економії). Але ж у комунікативному акті беруть участь дві сторони – „мовець” і „слухач”. І ставлення їх до економії мовленнєвих зусиль різне. Якщо для мовця вигідно користуватися короткими мовними одиницями, то для слухача це – не кращий варіант. Адже у цьому разі йому неодмінно доводиться стикатися з полісемією, витрачаючи зайві зусилля на пошуки відповідного значення. В загальнолітературній мові баланс сил – на боці мовця. Тому тут діє принцип найменших зусиль Ципфа.

Розходження спостережених даних з типовими для принципу Ципфа стосується зв'язку між довжиною та кількістю різних термінів у словнику й у тексті. В той же час залежність між лексичною, графічною довжиною термінів та частотою їх використання у текстах підлягає загальному закону: коротші терміни вживаються частіше (табл. 2).

Залежність між лексичною довжиною термінів та частотою вживання особливо чітко спостерігається при графічному зображенні (рис. 4).

Таблиця 2

Залежність частоти вживання терміна від його довжини

№ п/п	Кількість структурних компонентів, <i>L</i>	Середня частота	95-відсотковий довірчий інтервал		Кількість термінів
			Нижня границя	Верхня границя	
1	1	9,017	7,420	10,613	480
2	2	2,708	2,454	2,962	2372
3	3	1,706	1,586	1,826	1163
4	4	1,483	1,305	1,661	298
5	5	1,686	1,144	2,228	86
6	6	1,150	0,921	1,379	20
7	7	1,000	-	-	12
8	8	1,000	-	-	2

9	В середньому	3,014	2,783	3,245	4433
---	--------------	-------	-------	-------	------

На рис. 4 суцільна лінія відображає середню частоту вживання, короткі горизонтальні риси окреслюють межі довірчого інтервалу.

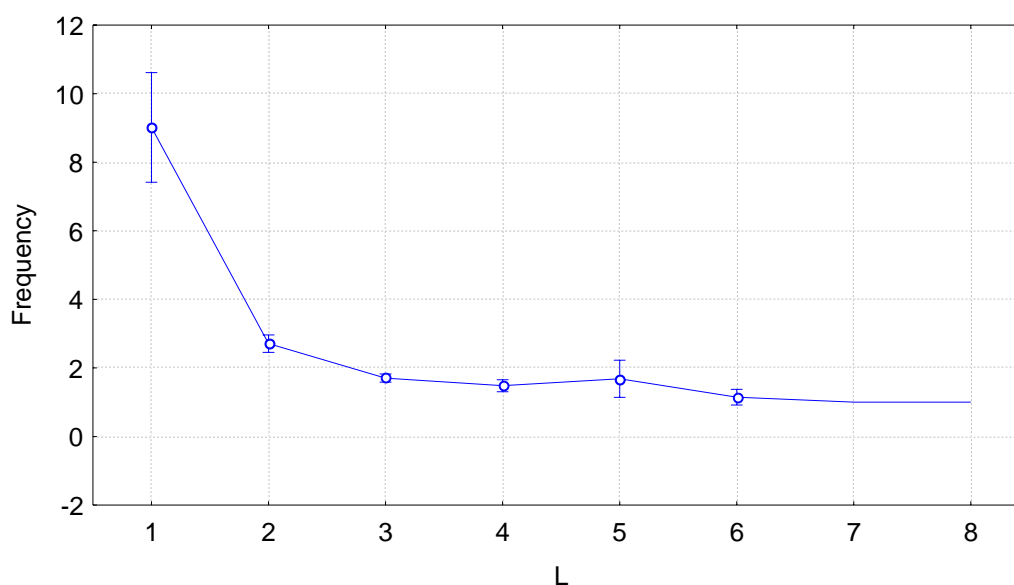


Рис. 4. Залежність частоти вживання терміна від його довжини

У термінології величезна кількість нових понять. Використання коротких мовних знаків на їх позначення спричиняє розвиток полісемії, активне використання семантичного термінотворення, зменшення ступеня вмотивованості термінів. Все це суттєво погіршує комунікативний клімат. Тому в термінологічній лексиці компроміс між поглядами мовця й слухача на економію зусиль, так само як і компроміс між різними, часто протилежними, вимогами до терміна, привів до того, що сформувалася лише відносно невелика кількість коротких, однокомпонентних, термінів на позначення базових понять. Основна маса спеціальних термінів налічує по два-три компоненти. Дослідження В.Д.Б'ялика виявили, що саме такі терміни є краще умотивовані [15]. Складні спеціальні поняття вимагають складніших мовних знаків, таких, що налічують більше трьох компонентів.

Перспективи подальших досліджень. Проведене дослідження дозволяє поглибити вивчення довжини **екстрагалузевих** термінів, які функціонують у англomовних наукових текстах зі штучного інтелекту,

матеріалознавства, аерокосмічних технологій, фінансово-економічної справи.

ЛІТЕРАТУРА

1. Д'яков А. С. Основи термінотворення / А. С. Д'яков, Т. Р. Кияк, З. Б. Куделько; – К. : КМ Academia, 2000. – 216 с.
2. Дерді Е. М. Словотвірні та структурно-семантичні характеристики англійських юридичних термінів : автореф. дис. на здобуття наук. ступеня канд. філол. наук : спец. 10.02.04 “Германські мови” / Емма Тіберіївна Дерді ; Київськ. нац. ун-т імені Тараса Шевченка. – К., 2003. – 21 с.
3. Гумовська І. М. Англійська юридична термінологія в економічних текстах: генезис, дериваційні та семантико-функціональні аспекти : автореф. дис. на здобуття наук. ступеня канд. філол. наук : спец. 10.02.04 “Германські мови” / Ірина Миколаївна Гумовська ; Львівськ. нац. ун-т імені Івана Франка. – Львів, 2000. – 19 с.
4. Скороходько Е. Ф. Термін у науковому тексті (до створення терміноцентричної теорії наукового дискурсу) : [монографія] / Едуард Федорович Скороходько. – К. : Логос, 2006. – 99 с.
5. Гринев С. В. Введение в терминоведение : [монография] / Сергей Васильевич Гринев. – М. : Московский лицей, 1993. – 310 с. : табл. – Библиогр. : с. 300–309.
6. Скороходько Е. Ф. Сучасна англійська термінологія : [навч. посіб.] / Едуард Федорович Скороходько. – К. : УІЛМ, 2002. – 76 с.
7. Ziegler A. Word Length in Romance Languages // *Quantitative Linguistics*. – 2000. – Vol. 7, № 1. – P. 65–68.
8. New, Boris et al. Reexamining the Word Length Effect in Visual Word Recognition // *Psychonomic Bulletin & Review*. – 2006. – Vol. 13, № 1. – P. 45–52.
9. Cowan N. List Composition and the Word Length Effect in Immediate Recall N. Cowan // *Psychonomic Bulletin & Review*. – 2003. – Vol. 10, № 1. – P. 74–79.
10. Antic Gordana. Word Length and Frequency Distributions in Different Text Genres / Antic Gordana // *Proceedings of the 29th Annual Conference [“From Data and Information Analysis to Knowledge Engineering”]*, (March 9–11 2005). – Magdeburg, 2005. – P. 18–26.
11. Grzybek P. Quantitative Text Typology. The Impact of Word Length / P. Grzybek // *Classification. – The Ubiquitous Challenge*. Springer, Heidelberg, 2005. – P. 53–64.
12. Скороходько Е. Ф. Сіткове моделювання лексики: лінгвістична інтерпретація параметрів семантичної складності // *Мовознавство : наук. теор. журнал / гол. ред. Віталій Григорович Складенко*. – К., 1995. – № 6. – С. 19–28.
13. Kelih E. Classification of Author and/or Genre? The Impact of Word Length / E. Kelih // *Classification. – The Ubiquitous Challenge*. Springer, Heidelberg, 2005. – P. 498–505.
14. Скороходько Э. Ф. Семантическая продуктивность и семантическая емкость слова в общеупотребительной и терминологической лексике: некоторые количественные характеристики / Эдуард Федорович Скороходько // *Научно-техническая информация. – Серия 2. – 1997. – № 2. – С. 1–13*.
15. Бялик В. Д. Структурно-семантические особенности и мотивированность научно-технических терминов в современном английском языке (на материале терминологии вычислительной техники) : автореф. дис. на соискание науч. степени канд. филол. наук / В. Д. Бялик ; Киевский гос. ун-т им. Т. Г. Шевченко. – К., 1986. – 25 с.

СПИСОК ДЖЕРЕЛ ІЛЮСТРАТИВНОГО МАТЕРІАЛУ

1. Bunke H. Recent Advances in Graph Matching / H. Bunke, B. T. Messmer // *International J. of Pattern Recognition and Artificial Intelligence*. – 1997. – Vol. 11, № 1. – P. 169–203.
2. Campanile L. F. Aerodynamic and aeroelastic amplification in adaptive belt-rib airfoils / Lucio Flavio Campanile, Stefan Anders // *Aerospace Science and Technology*. – 2005. – № 9. – P. 55–63.