

УДК 81'322.5'37:341.76(73)

ЗАСТОСУВАННЯ АВТОМАТИЗОВАНОГО КОНТЕНТ-АНАЛІЗУ ДЛЯ ДОСЛІДЖЕННЯ ТЕКСТОВИХ КАТЕГОРІЙ (НА МАТЕРІАЛІ ПРЕЗИДЕНТСЬКИХ МЕМОРАНДУМІВ БАРАКА ОБАМИ)

Денисюк Ю. І.

Національна академія сухопутних військ імені гетьмана Петра Сагайдачного

У статті досліджено текстові категорії президентських меморандумів Барака Обами з використанням комп'ютерної програми контент-аналізу текстів Tropes. Особливу увагу звернено на семантично-прагматичні чинники змістової організації текстів. З опорою на автоматизований контент-аналіз досліджуваних меморандумів обґрунтовано виокремлення модальностей характеру дії (manner), заперечення (negation), рівня напруженості (intensity), високого рівня переконаності (assertion) та низького рівня переконаності (doubt).

Ключові слова: програма контент-аналізу, лексико-семантична категорія тексту, прагматична категорія тексту, президентський меморандум.

Денисюк Ю. И. Применение автоматизированного контент-анализа для исследования текстовых категорий (на материале президентских меморандумов Барака Обамы). В статье исследуются текстовые категории президентских меморандумов Барака Обамы с использованием компьютерной программы контент-анализа текстов Tropes. Особое внимание уделено семантико-прагматическим факторам содержательной организации текстов. С опорой на автоматизированный контент-анализ исследуемых меморандумов обосновывается выделение модальностей характера действия (manner), отрицания (negation), уровня напряженности (intensity), высокого уровня уверенности (assertion) и низкого уровня уверенности (doubt).

Ключевые слова: программа контент-анализа, лексико-семантическая категория текста, прагматическая категория текста, президентский меморандум.

Denisjuk Yu. I. Application of automated content analysis for the study of text categories (based on the presidential memoranda of Barack Obama). The article deals with the analysis of textual modalities applying Tropes – a computer programme developed by Semantic-Knowledge and designed for semantic classification, keyword extraction and qualitative analysis of text. The research is based on Barack Obama's presidential memoranda (586 texts issued during his presidency in 2006 – 2016). It is claimed that textual modalities traced by the applied programme are correlated with their presentation in modal logic which deals with modal markers that resemble modal words in natural languages. Textual modalities as presented in this article do not coincide with textual categories as treated in modern text linguistics. Out of seven modalities chosen for text analysis by the developers of Tropes two are referred to semantic categories (time and place) and five – to pragmatic categories (manner, assertion, doubt, negation, intensity). Special attention is paid to the pragmatic categories as used by Barack Obama in his memoranda. It has been discovered that modalities of manner, negation and intensity are prevailing. Preliminary analysis of verbal means which represent these categories in presidential memoranda has been conducted.

Key words: content analysis program, lexico-semantic category of text, pragmatic category of text, presidential memorandum.

Постановка проблеми та обґрунтування актуальності її розгляду. Вчення про текстові категорії є сьогодні перспективним напрямом мовознавчих досліджень. Текст і його категорії мовознавці тлумачать неоднозначно, хоча формування теоретичних положень лінгвістики тексту почалось ще у 60-ті роки 20 століття. Сьогодні випрацювана низка різних підходів до виокремлення та трактування текстових категорій. Їх вивчення є ключовим питанням лінгвістики тексту [5].

Його розглядали такі дослідники у цій галузі, як В. Дресслер, Р. Богранд, М. Халлідей, А. Хілл, В. Бухбіндер, Є. Розанов, І. Гальперін, О. Селіванова. Більшість з них дотримуються поділу текстових категорій на формально-структурні й змістовні.

Однак Ф. Бацевич стверджує, що цей поділ є умовним, оскільки так зване «формальне» завжди

нерозривно пов'язане зі «змістовним» [1, 184]. Загалом у категорійній організації тексту Ф. Бацевич виділяє чотири типи категорій: 1) формально-структурні; 2) семантичні; 3) змістовно-семантичні; 4) прагматичні [1].

До першого типу відносяться категорії зчеплення, інтегративності, прогресії та стагнації. До другого типу – категорії темпоральності та локативності. Третій тип включає категорії «людина», «річ», «подія», «випадоку», «факт» та «оцінка». До прагматичних категорій (четвертий тип) Ф. Бацевич зараховує текстові категорії «точка зору» та «емпатія» й текстову модальність [1, 181–245].

Кожна з названих вище категорій привертає увагу дослідників і по-різному тлумачиться. Зокрема, поняття модальності, яке в мовознавстві має тривалу історію, по-новому висвітлюється з погляду її тлумачення як прагматичної категорії тексту. Дослідження є **актуальним** з огляду на зацікавленість мовознавців питаннями категорійної організації тек-

сту та питаннями виокремлення категорій (модальностей) тексту спеціальними автоматизованими програмами контент-аналізу.

Аналіз останніх досліджень і публікацій.

Сучасні дослідники тексту нерідко звертаються до автоматизованих методів його аналізу. Наприклад, зростає роль контент-аналізу в дослідженні текстів, породжених інститутами влади, які є комунікативними посередниками в соціальному регулюванні і які завдяки керівній ролі в суспільстві забезпечують соціальний примус [2].

Контент-аналіз – це найбільш усталений метод текстового аналізу серед емпіричних методів соціальних досліджень. [9, 596–692.] Як багатфункціональний метод аналізу тексту він може також прислужитися для виділення категорій тексту. Застосування контент-аналізу розпочалося у другій половині 1930-х років, передусім, для дослідження мас-медійного та міжнародно-політичного дискурсів [12, 254].

Сутність контент-аналізу полягає в тому, що кожному одиницю аналізу заковано та закріплено за однією чи більше категоріями. Будь-яке визначення обраних категорій повинно бути експліцитним, повним і адекватним. У процесі автоматизованого контент-аналізу застосовуються програми, які співвідносять одиниці аналізу з кодами [13]. Текстова категорія здебільшого постає як номінальна шкала: одиниця аналізу або належить або не належить до цієї категорії [8].

Загальна система категорій спрямована на операціоналізацію змінних величин, які можуть забезпечити опис конкретного об'єкта дослідження. Вперше перелік категорій для контент-аналізу уклали О. Холсті та Б. Берельсон. Вони запропонували типологію, яку можна використовувати як основу для проектування системи категорій: 1) предмет, тема (про що йде мова); 2) спрямованість (як розгортається тема); 3) норми (що слугує основою для класифікації та оцінки); 4) цінності (які установки, цілі та бажання виражаються); 5) засоби (які засоби використовуються для досягнення поставлених цілей); 6) особливості (які особливості використовуються при описі людей); 7) актори (хто ініціює певні дії, а хто їх виконує); 8) авторство (від імені кого зроблені твердження); 9) джерело (хто здійснює комунікацію); 10) конфлікт (яка причина будь-якого конфлікту, хто учасники, наскільки серйозний конфлікт); 11) результат (чи розв'язок конфлікту є позитивним, трагічним чи невизначеним); 12) час (коли все відбувається); 13) форма або вид комунікації (який канал комунікації використовується); 14) форма твердження (які граматичні і синтаксичні форми можна виявити); 15) методи (які риторичні прийоми чи методи пропаганди застосовуються) [6, 96]. Цей список показує, що контент-аналіз не обмежується лише експліцитним вмістом тексту – його метою є знайти відповіді на ті та інші питання з випрацюванням схеми текстових категорій.

Мета пропонованої статті – проаналізувати текстові категорії президентських меморандумів (далі ПМ) Барака Обами за 2009–2016 рр. за допомогою програми автоматизованого контент-аналізу Tropes. Це програмне забезпечення розробила компанія Semantic-Knowledge у 1994–2010 рр. [7]. Tropes забезпечує контент-аналіз англійських текстів на засадах семантичної класифікації. Аналіз тексту здійснюється за такими параметрами: стиль тексту, референтні поля, референти, сценарій, зв'язки референтних полів, частотні категорії слів та всі категорії слів.

Для досягнення поставленої мети потрібно розв'язати такі **завдання**:

- виявити найчастотніші текстові категорії цього мовленнєвого жанру;
- систематизувати мовні засоби, які вербалізують кожен з категорій.

Виклад основного матеріалу дослідження.

ПМ належать до адміністративних наказів в складі президентського дискурсу і є важливим складником жанрової палітри цього дискурсу. ПМ мають чітку структуру (дет. див. [3]). За час президентства (2009–2016 рр.) Барак Обама видав 586 ПМ (табл. 1).

У програмі Tropes у підрозділі «Текстові категорії (модальності)» виділено сім категорій, і програма підраховує відсоткові та кількісні показники реалізації цих модальностей (рис. 1).

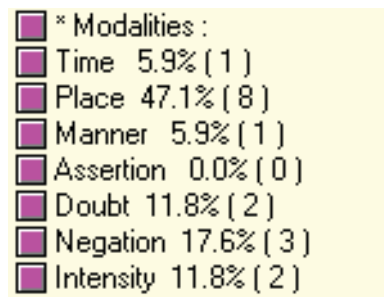


Рис. 1. Приклад представлення текстових категорій (модальностей) у програмі Tropes

Модальності в контент-аналізі співвідносяться з їх тлумаченням у модальній логіці. Остання розглядає логічні системи, у яких використовуються модальні оператори, подібні за змістом до модальних дієслів природної мови: «можна», «потрібно», «можливо», «неможливо», «необхідно», «обов'язково», «дозволено», «заборонено», «впевнений», «було», «буде» та інші [4].

Залежно від того, яка модальність надає висловлюванню нового смислу, оцінює те, що стверджується або заперечується, логіки розрізняють типи модальностей: 1) алетичні («необхідно», «можливо», «насправді», «випадково»); 2) деонтичні («обов'язково», «дозволено», «заборонено»); 3) епістемічні («знає», «вірує», «сумнівається», «відомо»,

Таблиця 1

Кількість президентських меморандумів Барака Обами, виданих у 2009-2016 роках

2009	2010	2011	2012	2013	2014	2015	2016
66	70	84	83	50	44	70	119

«невідомо», «переконаний»); 4) часові («було», «є», «буде») тощо.

Кожен тип модальності визначають з погляду об'єктивних (фактичних) і логічних детермінантів. Алетична модальність, зокрема, детермінується з погляду об'єктивних законів природи, суспільства й логічних законів; деонтична – з погляду прийнятих у суспільстві правових і моральних норм; епістемічна – з погляду закономірностей пізнання об'єктивного світу, часова – з погляду часових характеристик того, що відбувається у світі [4].

Таким чином, тлумачення модальностей у логіці (і, відповідно, у системах контент-аналізу) виходить за межі тлумачення модальності як прагматичної категорії тексту в мовознавстві, а охоплює також інші типи категорій. Темпоральність та локативність (рис. 1) є семантичними категоріями тексту, які в цьому дослідженні не розглядаються. Інші категорії, зокрема: 1) характер дії (*manner*); 2) високий рівень переконаності (*assertion*); 3) низький рівень переконаності (*doubt*); 4) заперечення (*negation*); 5) рівень напруженості (*intensity*) – розглядаються як різновиди прагматичних текстових категорій. Зосередимо увагу на реалізації цих прагматичних категорій у ПМ Барака Обами. Застосування програми Tropes дозволило встановити, які текстові категорії трапляються частіше і якими мовними засобами вербалізується кожна з них.

Кількісні показники реалізації п'яти категорій, за якими програма Tropes здійснює аналіз текстів, представлено в таблиці 2.

Аналіз цих показників засвідчує переважання модальності характеру дії (*manner*), заперечення (*negation*) та рівня напруженості дії (*intensity*). Найчастотнішою текстовою категорією є категорія характеру дії (*manner*) – 377 випадків вживання, другою за частотністю є категорія заперечення (*negation*) – 288 випадків, а на третьому місці – напруженість дії (*intensity*), що становить 269 випадків. Категорії високого (*assertion*) та низького (*doubt*) рівня переконаності трапляються значно рідше (відповідно 60 та 28 випадків).

Перелічені модальності є визначальними для жанру ПМ, оскільки комунікативну спрямованість текстів цього жанру експліковано мовленнєвими актами, що виражають настанови, розпорядження та несхвалення (дет. див [3]).

Програма автоматизованого контент-аналізу також забезпечує та значно спрощує виявлення тих мовних засобів, які вербалізують кожен з текстових категорій. Аналіз досліджуваних меморандумів за роками дозволив систематизувати ці мовні засоби (до прикладу, див. таблицю 3).

У різних текстах ПМ текстові категорії можуть бути представлені різними мовними одиницями. До прикладу, розглянемо меморандум, написаний Бараком Обамою 10 травня 2013 року [11]. У цьому меморандумі йдеться про покращення системи оплати праці. Президент накреслює план конструктивних змін і дає чіткі вказівки урядовим структурам та відомствам, спрямовані на організацію їхньої роботи для досягнення поставлених завдань. У цьому ПМ реалізуються категорії характеру дії (*manner*), низького рівня переконаності (*doubt*), заперечення (*negation*) та рівня інтенсивності дії (*intensity*). Категорію характеру дії (*manner*) представлено маркерами: *broadly, similarly, how, part-time*:

In order to further understand how the practices of executive departments and agencies (agencies) affect the compensation of similarly situated men and women, and to promote gender pay equality in the Federal Government and more broadly, I hereby direct the following actions, pursuant to the authority vested in me by the Constitution and the laws of the United States) [11].

Категорія низького рівня переконаності (*doubt*) вербалізується через маркер: *almost: Almost 50 years ago, when President John F. Kennedy signed the Equal Pay Act of 1963, women were paid 59 cents for every dollar paid to men) [11].*

Категорії заперечення (*negation*) позначено маркерами: *not, nothing: Nothing in this memorandum shall be construed to impair or otherwise affect, This memorandum is not intended to, and does not, create any right or benefit, substantive or procedural, enforceable at law or in equity by any party against the United States, its departments, agencies, or entities, its officers, employees, or agents, or any other person) [11].*

Категорія рівня інтенсивності дії (*intensity*) виражається словами: *nearly, more, particular: At the same time, nearly two-thirds of women are breadwinners or co-breadwinners for their families, all agency-specific policies and practices for evaluating individuals regarding promotions, particularly individuals who work part-time schedules (for example,*

Таблиця 2

Кількісні показники реалізації прагматичних категорій у президентських меморандумах Барака Обами

Рік	Кількість випадків реалізації категорії				
	Manner	Assertion	Doubt	Negation	Intensity
2009	44	1	4	50	34
2010	42	9	3	46	36
2011	55	8	8	44	41
2012	45	8	2	28	27
2013	27	4	6	27	21
2014	30	4	3	28	20
2015	53	11	1	24	41
2016	81	12	4	41	49

Таблиця 3

Вербалізація текстових категорій у ПМ Барака Обама за 2013 рік

	Текстові категорії				
	Manner	assertion	doubt	negation	intensity
Вербальні засоби	as follows, precisely, readily, how, optimally, appropriately, such, economically feasible, technically, as required, publicly, separately, globally, internationally, directly, at large, broadly, adversely, deeply, efficiently, currently, reasonably, potentially, adversely, likewise, publicly, repeatedly, in collaboration, constructively, carefully, lastly, additionally, primarily, habitually, continuously, voluntarily, accordingly, similarly, part-time, lawfully, electronically, as needed, sincerely, jointly, in particular, significantly, conscientiously, together, equally, strategically, carefully, reasonably, in the air, demonstrably, as required, especially, bilaterally, clearly	effectively, exclusively	almost, approximately	not, no, nothing, no longer, negatively	also, particularly, well, more than, better, so, fully, substantially, even more, most, strongly, seriously, nearly, all, best, at last, only, each, at all, less

those who serve as caregivers to children or other family members) [11].

Порівняльний аналіз мовних засобів за 8 років президентства засвідчує повторюваність маркерів, які характерні для кожної модальності жанру ПМ.

Висновки та перспективи подальших досліджень. Підсумовуючи, зазначимо, що застосування програми автоматизованого контент-аналізу забезпечує виявлення в текстах семантико-прагматичних чинників їхньої змістової організації. Модальності тексту, пошук яких забезпечує програма, співвідносні, однак

не збігаються з текстовими категоріями, які розглядають сьогодні у лінгвістиці тексту. Застосування програми контент-аналізу Tropes для розгляду ПМ Барака Обама за 2009–2016 дозволило зробити висновок про переважання в цих текстах модальностей характеру дії (*manner*), заперечення (*negation*) та рівня напруженості дії (*intensity*), які є виразниками суб'єктивної модальності в досліджених меморандумах. Перспективним видається докладний аналіз засобів вербалізації модальностей та встановлення їхньої ролі в досягненні комунікативних цілей президента.

ЛІТЕРАТУРА

1. Бацевич Ф. Лінгвістика тексту / Флорій Бацевич, Ірина Кочан. – Львів : ЛНУ імені Івана Франка, 2016. – 316 с.
2. Волков А. Г. Прагматика политического дискурса / А. Г. Волков. – Запорожье : Просвіта, 2003. – 112 с.
3. Денисюк Ю. І. Типи президентських меморандумів Барака Обама : Структурні особливості / Ю. І. Денисюк // Науковий вісник Дрогобицького державного педагогічного університету імені Івана Франка. Серія: "Філологічні науки" (мовознавство) : [зб. наук. праць]. – Дрогобич, 2016. – № 5. – Т. 1. – С. 108–111.
4. Карамішева Н. В. Логіка. Пізнання. Евристика : [посіб. для студентів та аспірантів] / Н. В. Карамішева. – Л. : Астролябія, 2002. – 352 с.
5. Николаева Т. М. Лингвистика текста / Т. М. Николаева // Лингвистический энциклопедический словарь. – М. : Сов. энциклопедия, 1990. – 682 с.
6. Титчер С. Методы анализа текста и дискурса / Стефан Титчер. – Х. : Изд-во Гуманитарный Центр, 2009. – 356 с.
7. Content analysis programme [Електронний ресурс]. – Режим доступу : www.semantic-knowledge.com
8. Herkner W. Inhaltsanalyse / Werner Herkner // [ed. by Juergen von Koolwijk and Maria Wicken-Mayser Techniken der empirischen Sozialforschung]. – Muenchen : Oldenbourg, 1974. – Vol. 3. – P. 158–191.
9. Holsti O. Content analysis / Ole R Holsti / [ed. by G. Lindzey and E. Aronson] // The Handbook of Social Psychology, 2nd edn., 1968. – Vol. 2. – P. 596–692.
10. Merten K. Inhaltsanalyse. Einfuehrung in Theorie, Methode und Praxis / Klaus Merten // Opladen : Westdeutscher Verlag, 1983. – P. 3–20.
11. Presidential Documents Guide [Електронний ресурс]. – Режим доступу: <https://obamawhitehouse.archives.gov/the-press-office/2013/05/10/presidential-memorandum-advancing-pay-equality-federal-government-and-le>
12. Silbermann A. Systematische Inhaltanalyse / Alphons Silbermann / [ed. by Rene Koenig Handbuch der empirischen Sozialforschung]. – Stuttgart : Komplexe Forschungsansätze, 1974. – Vol. 4. – P. 253–339.
13. Weitzman E. Computer programs for Qualitative Data Analysis / Eben Weitzman and Miles Mathew // Thousand Oaks : Sage. Publications, 1995. – 371 p.