

Григорій Рудий

## Методика аналізу газетної періодики України (1917–1920 рр.)

За підрахунками автора, в 1917-1920 рр. на теренах України видавалося 1152 газети. Аналіз газетної періодики часів Української революції як історичного джерела дав змогу визначити якісні та кількісні характеристики тих явищ і процесів, що були предметом газетних публікацій.

Головним чинником, який спричинив звернення до кількісного аналізу, є потреба у повнішому, глибшому й точному розкритті сутності реальних, політичних, економічних та соціально-культурних процесів, що відбулися в матеріалах преси. При вивченні великих інформаційних масивів газетної періодики України в досліджуваній період виникає необхідність в точнішому, конкретному аналізі кількісних характеристик, зокрема виявлення інформаційних і ціннісних елементів у змісті повідомлень, вивчення принципів відбирання та розміщення матеріалів, пошуки найпопулярніших тем і жанрів. При цьому слід мати на увазі: ніякого суто самостійного описового чи кількісного аналізу не існує. Може бути лише єдність у зазначеному співвідношенні.

Як свідчить практика, найефективнішим при опрацюванні газетних матеріалів є контент-аналіз<sup>1</sup> метод кількісного вивчення змісту соціальної інформації<sup>2</sup>. Від традиційних методів аналізу текстової інформації, застосовуваних у джерелознавстві, літературознавстві та інших гуманітарних науках, контентний аналіз відрізняється значнішою точністю, системністю, використанням кількісних показників. Як будь-яка кількісна процедура, контент-аналіз замінює інтуїтивне сприйняття конкретнішою інформацією, дає змогу порівнювати групи фактів з кількісно визначеними символічними даними: словами, висловами, поняттями тощо. На стадії простого опису цей аналіз швидше породжує точнішу міру, чим поліпшує глобально або інтуїтивно сприйняття.

Рудий Григорій Якович, к. і. н., науковий співробітник-консультант Науково-організаційного відділу Президії НАН України.

Цінність контентного аналізу залежить від якості концептуалізації, проведеної дослідником до аналізу, від тієї точності, з якої поняття переводяться в перемінні, від схеми аналізу та категорій. Іншими словами, для застосування контентного аналізу необхідно, по-перше, щоб поставлені питання, на які треба відповісти на підставі дослідження, виявили цікаву гіпотезу і, по-друге, щоб питання відповідали цьому матеріалу.

Ми поділяємо думку фахівців, котрі вважають, що цінність контент-аналізу залежить від поставлених у дослідженні гіпотез і категорій, висловлених у них<sup>3</sup>.

Техніка кількісного аналізу знаходить дедалі ширше застосування в історичних дослідженнях. Цей метод є особливо ефективним при вивченні проблем освіти, науки, культури, що відображаються в пресі. Тією мірою, якою культурно-освітні процеси піддаються ідеї монтажу, контент-аналіз є доречним і ефективним, бо дає змогу дослідникові зібрати воедино окремі смислові цінності й символічні елементи і розглянути пов-ноцінну картину під певним кутом зору.

Квантифікаторний аналіз аналогічних текстів спрямований, насамперед, на характеристику змістовних і структурних блоків<sup>4</sup>. З цією метою виділяються одиниці обліку, смислові або якісні одиниці аналізу. Вважається, що одиницею інформації є та її кількість, яка дає змогу зробити вибір між двома рівновірогідними джерелами.

Досліджуючи газетні матеріали, можна виділити такі одиниці обліку:

- а) поняття, висловлені в словах і окремих термінах. Наприклад, за частотою вживання понять, пов'язаних з діяльністю товариств «Просвіта», можна визначити, якою мірою газетна періодика зорієнтована на розвиток просвітянського руху в Україні;
- б) тема — це більший фрагмент, що відповідає ідеї, яка охоплює одну з категорій. Тема — елемент менш точний, ніж слово, оскільки



вона може займати як два рядки, так і кілька сторінок;

- в) назви державних установ, політичних партій, громадських організацій, прізвища державних, політичних, культурних діячів. Дані характеристики свідчать про вплив окремих особистостей або установ, партій, громадських інституцій на формування громадської думки;
- г) предмет. Цей, безперечно, найвживаніший термін у широкому тлумаченні слова означає зміст газети цілком. Він може бути синонімом одиниці обліку, тобто всього того, що піддається виміру. Слово, тема, назва, прізвища, вживані як одиниці реєстрації, розглядаються просто як предмети. У деяких випадках за одиницю обліку беруть фізичну площу газетних матеріалів, заповнену смисловими одиницями: кількість рядків, абзаців, квадратних сантиметрів, знаків, колонок.

У зв'язку з цим постає питання: чи вимірює одиниця аналізу саме те, що вона мала б вимірювати?

Аналіз достовірний, якщо квантифікований опис змісту є значним для досліджуваної проблеми і точно відбиває реальні факти.

Незважаючи на досить поширене використання цього методу, в джерелознавчій практиці ще не сформульовано стале визначення терміна «контент-аналіз», що мало б охоплювати його найсуттєвіші методологічні та функціональні ознаки. Зокрема, деякі фахівці намагаються розглядати контент-аналіз як метод формалізованого спостереження і статичних процедур при аналізі масових сукупностей текстів<sup>5</sup>.

Актуальним, на нашу думку, є питання про надійність, достовірність і валідність (обґрунтованість) контентного методу. Системність, послідовність, об'єктивність, точність даних, їх суворі перевірки характеризують цей метод і гарантують надійність і достовірність отриманих результатів. Також ми поділяємо думку, що контент-аналіз валідний, бо дозволяє досліднику вимірювати саме ту властивість досліджуваного ним об'єкта, яку й вимагалось з'ясувати<sup>6</sup>. Як свідчить практика, контентний аналіз передбачає кількісний і якісний аналіз матеріалів преси. Суть цього комбінованого методу полягає в тому, що, по-перше, необхідно окреслити коло лише тих символів, інформантів, які конкретизують уже сформульовані категорії, по-друге, дають змогу визначити частотність інформаційних ознак, по-третє, створюють необхідні умови для оцінки якісного стану групи джерел з зазначеної проблематики<sup>7</sup>.

Враховуючи досвід як у галузі теоретико-методологічної розробки, так і джерелознавчої прак-

тики, ми вважаємо за доцільне розглядати контент-аналіз як точний науковий метод, що передбачає систематичну, послідовну, надійну, валідну фіксацію певних елементів змісту матеріалів преси з наступним кількісним опрацюванням отриманих даних.

Саме з цих позицій нами опрацьовано матеріали газет вказаного періоду. Для дослідження газетних джерел доцільно визначити фізичний обсяг матеріалів за окремими тематичними напрямками для вирахування їхньої питомої ваги серед комплексу інших джерел. Для цього здійснено аналіз джерела з метою з'ясування, насамперед, розвитку національної культури у зазначений період.

Отже, постає необхідність вибору категорій аналізу. Це питання є ключовим у контент-аналізі. Категорії утворюють зв'язок між метою дослідження і його результатами. У фаховій літературі поширюється думка, що контентний аналіз вартісний остільки, оскільки варті уваги його категорії<sup>8</sup>, а для визначення категорій слід дотримуватися таких правил:

- а) вони мають бути насамперед вичерпними, тобто придатними для розгляду всього змісту;
- б) категорії повинні мати ознаки винятковості: одні й ті самі елементи не повинні належати до кількох категорій;
- в) категорії мають бути об'єктивними, їх характеристики повинні бути зрозумілими для різних кодувальників, щоб вони могли класифікувати різні елементи змісту з допомогою однакових категорій;
- г) категорії мають бути доречними, відповідаючи переслідуючій меті і досліджуваному змісту. Якщо категорії передбачаються заздалегідь, то виникає небезпека пожертвувати їхньою доречністю<sup>9</sup>.

Щоб скласти загальну основу для вибору категорій контент-аналізу текстів періодичної преси, на нашу думку, насамперед треба звернутися до визначення соціальних функцій, які виконує періодика в суспільстві. Матеріали газет—це об'єктивація системи взаємозв'язку по лінії комунікатор — читач. Інформація виникає там, де є не тільки джерело, а й об'єкт із достатньо високим рівнем організації, виникає в результаті їх взаємодії<sup>10</sup>. Таке розуміння допускає існування в тексті (текстовій реальності) всього того, що існує в реальності поза текстом. Отже, принцип вибору категорій можна сформулювати так: схема категорій, що задає структуру фіксуемого в тексті об'єкта має в загальних рисах відтворити його модель для позатекстової реальності. Таким чином, категорії є значними рубриками, згідно з якими кваліфікується і квантифікується зміст. Виходячи з цих методологічних засад, нами складено відповідні тематичні блоки за їх складовими категоріями:



**Блок № 1. Відображення пресою шляхів побудови національної школи і науки**

**Категорії:**

- 1) діяльність державних органів і громадських організацій при створенні Української Академії наук у Києві;
- 2) розвиток УАН;
- 3) діяльність Українського наукового товариства, наукових товариств;
- 4) створення і діяльність українських державних університетів та інших вищих навчальних закладів, українізація російськомовних вишів;
- 5) роль освітянських державних органів у становленні національної початкової та середньої школи;
- 6) створення і діяльність шкіл для дорослих;
- 7) заснування і діяльність українських початкових шкіл;
- 8) створення і діяльність гімназій та середніх шкіл;
- 9) підготовка учительських кадрів.

**Блок № 2. Відродження української мови, літератури і мистецтва в дзеркалі преси**

**Категорії:**

- 10) відродження і розвиток української мови;
- 11) розвиток української літератури;
- 12) відродження традицій і звичаїв минулого;
- 13) діяльність державних культурно-освітніх установ;
- 14) заснування і діяльність Української Академії мистецтв;
- 15) створення і діяльність товариств «Просвіта»;
- 16) відкриття і діяльність Національної бібліотеки Академії наук у Києві, українських державних бібліотек;
- 17) розвиток національних професійних театрів, кінематографа;
- 18) розвиток музичного мистецтва;
- 19) організація та підтримка художньої самодіяльності.

**Блок № 3. Відображення газетною періодикою процесу становлення державної системи охорони пам'яток історії та культури**

**Категорії:**

- 20) створення та діяльність державних органів по охороні пам'яток старовини й мистецтва;
- 21) обстеження і реставрація пам'яток історії та культури;
- 22) реконструкція і створення нових музеїв, пам'ятників.

**Блок № 4. Висвітлення у пресі України розвитку видавничої справи**

**Категорії:**

- 23) заснування, діяльність і закриття україн-

ських газет і журналів;

- 24) видання наукової і науково-популярної літератури;
- 25) випуск суспільно-політичної та історичної літератури;
- 26) видання художньої літератури;
- 27) випуск підручників, навчальної літератури;
- 28) створення і діяльність українських видавництв.

**Блок № 5. Масштаб проблеми**

**Категорії:**

- 1) події всеукраїнського значення;
- 2) проблеми місцевого значення.

**Блок № 6. Жанри періодичних видань**

**Категорії:**

- 1) допис;
- 2) кореспонденція;
- 3) репортаж;
- 4) хроніка;
- 5) стаття, рецензія;
- 6) інтерв'ю;
- 7) нарис;
- 8) фейлетон;
- 9) оповідання;
- 10) вірш;
- 11) лист.

**Блок № 7. Авторський склад**

**Категорії:**

- 1) працівники редакції;
- 2) державні, громадські, політичні, військові діячі;
- 3) учені, працівники освіти, культури;
- 4) журналісти, письменники;
- 5) інші категорії дописувачів.

У результаті дослідницьких процедур розроблено спеціальну облікову картку, що є основою для складання програми для обчислення матеріалу на ЕОМ.

Для машинної обробки всі дані облікової картки відповідним чином кодуються: у графі за № 5 міститься фізичний обсяг газетного джерела (у см<sup>2</sup>), а за № 6 — інформаційний, що охоплює кількість символів-інформантів з експертною оцінкою. Запропонована облікова картка відображає типологічну характеристику однієї публікації. При аналізі газетних матеріалів зазначеного періоду заповнено кілька тисяч таких карток. Зрозуміло, що даний масив формалізованої інформації може бути опрацьований тільки за допомогою комп'ютерної техніки. Одержані за розробленою схемою результати склали інформаційно-формалізовану основу для джерелознавчого аналізу даних, що містить поняття «видання». Виходячи з цього, нами розроблено програму, яка дає



змогу з'ясувати причинно-наслідкову залежність та перевірити гіпотези щодо бази даних.

Кожна гіпотеза  $H$ , що перевіряється, складається з двох частин: <умова> і <наслідок>. Позначимо збіг значень <умови> та полів першого запису бази даних як подію  $A$ . Незбіг значень хоч би й одного поля—це протилежна подія  $\bar{A}$ . Відповідно збіг значення <наслідок> із значенням досліджуваного поля з бази даних — подія  $B$ , незбіг —  $\bar{B}$ . Усіх можливих варіантів комбінацій чотири:  $(A, B)$ ,  $(\bar{A}, B)$ ,  $(A, \bar{B})$ ,  $(\bar{A}, \bar{B})$ . До перевірки гіпотези мають відношення записи, де відбувається подія  $A$ . Якщо записів із комбінацією подій  $(A, B)$  було  $N_1$ , а з подіями  $(A, \bar{B})$  —  $N_2$ , то достовірність гіпотези  $H$  підраховується за формулою:

$$G(H) = \frac{N_1}{N_1 + N_2} * 100\%$$

Це фактично є статистичною ймовірністю, оскільки вираховується відносна частота появи комбінацій  $(A, B)$ , а при даному обсязі вибірки вже може йтися про наближення її до достовірності.

Наприклад, потрібно вирахувати ймовірність того, що стаття, написана працівником культури, торкатиметься проблем охорони пам'яток старовини та культури. Умовною гіпотезою тут є:

Поле «жанр»: 5 (що відповідає статті); поле «автор»: 3 (працівник культури); наслідок — 20 (проблема охорони пам'яток історії та культури) у полі № 3 (тематичний розділ).

Зазначимо: відповідність чисел і значення (наприклад, 5 — це стаття) є «внутрішньою справою» програми, що опрацьовує базу даних, оскільки поля записів є саме числовими.

Якщо не задати умов гіпотези, а вказати лише наслідки, то можна знайти частку тих чи інших значень щодо загальної кількості записів. Наприклад, комбінація браку умов та значення 2 (проблема місцевого значення) в останньому полі дає змогу знайти відсоток публікацій, присвячених подіям місцевого значення стосовно всієї бази даних. Оскільки контент-аналіз становить собою сукупність аналітичних і синтетичних процедур, метою яких є з'ясування найсуттєвіших якостей, головних закономірностей та особливостей виникнення і функціонування об'єктів і досліджуваних явищ, то виникає необхідність визначити взаємозв'язок між ними. До методів аналізу взаємозв'язку належить аналіз дисперсійний, аналіз якісний регресивний, аналіз латентно-структурний, аналіз факторний, кореляційний аналіз та ін. У межах нашого дослідження вбачається доцільним обмежитися саме кореляційним аналізом. Кореляція — функціональна залежність між двома перемінними величинами, яка характеризується тим, що кожному значенню однієї з них

відповідає цілком певне значення другої. Якщо певному значенню однієї величини відповідає комплекс значень другої, йдеться про кореляційну залежність. При зміні однієї величини змінюється комплекс значень другої, який може бути схарактеризований своєю середньою. Якщо збільшенню однієї величини відповідає збільшення відповідних середніх другої, йдеться про пряму або позитивну кореляцію (в протилежному випадку — про обернену або негативну)<sup>11</sup>. Наочним прикладом є зв'язок між автором і тематикою, тематикою і жанром, жанром і автором, жанром і масштабом проблеми. Внаслідок складності, багатоінформативності газетної періодики зв'язки між інформаційними перемінними практично завжди кореляційні. У рамках нашого дослідження ми обмежуємо кореляційний аналіз з'ясуванням лінійної взаємодії між жанром і автором та автором і тематикою. Кількісно ці зв'язки з'ясовуються з допомогою коефіцієнта кореляції. Коефіцієнт кореляції — міра щільності. Коли кожному значенню однієї ознаки відповідають різні, але близькі значення іншої ознаки, тобто тісно розміщені біля своєї середньої величини, — зв'язок щільніший<sup>12</sup>.

Виходячи з цього, коефіцієнт кореляції між зазначеними величинами вираховується за такою формулою:

$$r_{xy} = \frac{K_{xy}}{V_x V_y}$$

Тут  $V_x$  та  $V_y$  — середньостатистичні відхилення,  $V_x = \sqrt{D_x}$ ,  $V_y = \sqrt{D_y}$ ;  $D_x$  та  $D_y$  — це дисперсії величин. Візьмемо незміщені оцінки:

$$D_x = \frac{1}{n-1} * \sum_{i=1}^n (X_i - M_x)^2 \quad D_y = \frac{1}{n-1} * \sum_{i=1}^n (Y_i - M_y)^2$$

$K_{xy}$  — це кореляційний момент:

$$K_{xy} = \frac{1}{n-1} * \sum_{i=1}^n (X_i - M_x)(Y_i - M_y)$$

Це також незміщена оцінка. За математичне сподіване  $M_x$  та  $M_y$  взято статистичне математичне сподіване (середнє значення):

$$M_x = \frac{1}{n} * \sum_{i=1}^n x_i, \quad M_y = \frac{1}{n} * \sum_{i=1}^n y_i$$

Крім цього, контент-аналіз української преси досліджуваного періоду дав змогу визначити кількісні характеристики (параметри) тих явищ і процесів, що були предметом газетних публікацій. В основу аналізу жанрів газетної періодики нами покладено сюжетні категорії, позначені в інформаційному блоці № 6. З цією метою використано



такі форми компактного викладу фактичного матеріалу, як графіки та гістограми. Підрахунки проведено за формулою:

$$P_i = \frac{p_i}{n} * 100\%,$$

де  $n$  — загальна кількість записів у базі даних.

За аналогічною формулою проведено підрахунки для побудови гістограм з середнього інформаційного обсягу для кожного жанру, а також по середньому відношенню інформаційного обсягу до фізичного. Вибудовані діаграми з тематики газетних публікацій окремо по кожному тематичному блоку. Зроблено підрахунки динаміки газетних матеріалів щодо масштабу проблеми.

З метою якісної оцінки змісту великих масивів текстів з актуальних проблем національної культури вбачається доцільним скласти методичний інструментарій для виявлення характерних символів-інформантів. Інструмент смислового простору об'єднує 27 понять: *наукові товариства; Академія наук; українізація; національна освіта; Генеральне секретарство освітніх справ; Міністерство народної освіти; Міністерство народної освіти та мистецтва; Наркомат освіти УСРР; університет, гімназія, середня школа, початкова школа; відродження мови, друкованого слова, українські видавництва; державна мова; національна культура; Департамент мистецтв; Головне управління мистецтв та національної культури; Національна бібліотека; товариства «Просвіта»; Українська Академія мистецтв; Державний театр; Національна опера; Національна галерея мистецтв; Українська державна капела; історія, традиції, минуле; пам'ятки старовини і мистецтва; церковна старовина; Центральний комітет охорони пам'яток старовини і мистецтва України; Український національний музей; Український національний хор.*

Як засвідчує джерелознавча практика, взаємовідносини між поняттями, категоріями і темами досить складні. Принаймні вони можуть бути постійними або перемінними між наборами символів-інформантів, між наборами категорій і темами. Найчастіше трапляється тип взаємозв'язку, який передбачає, що теми, в свою чергу, бувають операционовані одними й тими самими символами-інформантами<sup>13</sup>. Таким чином, результати інтерпретації символів як репрезентантів оцінюючих категорій зумовлюють інтерпретацію символів текстів як репрезентантів оцінювальних орієнтацій. Поряд з цим вбачається важливим виявлення критеріїв для виміру інформаційної насиченості джерел, їх репрезентативності. Для розв'язання цього завдання ми використали метод експертних оцінок. Він досить докладно висвітлюється в історичній літературі<sup>14</sup>. При вивченні

значних масивів джерел, що відбивають велике число об'єктів, виникає необхідність визначити критерії вимірювання. Це, на нашу думку, має неабияку вагу для якісної оцінки джерела.

Головним фактором, котрий викликає необхідність звернення до методу експертних оцінок, є потреба в більш глибокому й точному розкритті сутності явищ і процесів, які відбулися в матеріалах газетної періодики. А експертне оцінювання передбачає розв'язання таких завдань:

1) домовленість про те, що вибрані для аналізу якості і критерії оцінювання відповідають поставленим цілям;

2) побудова моделей розв'язань і вимірів, необхідних для оцінки даних;

3) вироблення єдиної підстави вибору конкретних варіантів розв'язань<sup>15</sup>.

У межах кожного дослідження ми з допомогою експертних оцінок конкретизували явища, пов'язані з інформаційною насиченістю газетних матеріалів.

Одержані результати, що розкривають внутрішній зв'язок між масивами джерел, створили можливість дати кількісну і якісну характеристику матеріалів преси періоду Української революції.

<sup>1</sup> Количественные методы в исторических исследованиях. — М., 1984. — С. 18.

<sup>2</sup> Краткий словарь по социологии. — М., 1989. — С. 122.

<sup>3</sup> Пэнто Р., Гравитц М. Методы социальных наук. — М., 1972. — С. 351.

<sup>4</sup> Хвостова К. В. Контент-анализ в исследованиях по истории культуры // Одиссей: Человек в истории. — М., 1989. — С. 136.

<sup>5</sup> Дергачева Л. Д. Периодическая печать как массовый источник // Массовые источники по истории советского рабочего класса периода развитого социализма. — М., 1969. — С. 159.

<sup>6</sup> Социологический справочник. — К., 1990. — С. 195.

<sup>7</sup> Ганжуров Ю. С. Преса України другої половини 50-х — початку 60-х рр. як об'єкт кліометричних досліджень у сучасному джерелознавстві // Укр. іст. журн. — 1991. — № 12. — С. 41—42.

<sup>8</sup> Пэнто Р., Гравитц М. Методы социальных наук. — С. 338.

<sup>9</sup> Там же.

<sup>10</sup> Семенюк Э. П. Информационный подход к познанию действительности. — К., 1988. — С. 45.

<sup>11</sup> Социологический справочник. — С. 215—216.

<sup>12</sup> Там само. — С. 217.

<sup>13</sup> Костенко Н. В. Ценности и символы в массовой коммуникации. — К., 1993. — С. 79.

<sup>14</sup> Количественные методы в исторических исследованиях. — С. 30—31.

<sup>15</sup> Экспертные оценки в социологических исследованиях. — К., 1990. — С. 18.