

СТАТИСТИЧНІ ХАРАКТЕРИСТИКИ СПОЛУЧУВАНОСТІ ЛЕКСИКО-СЕМАНТИЧНИХ ГРУП ПРИКМЕТНИКІВ З ІМЕННИКАМИ В АНГЛІЙСЬКІЙ ТА АМЕРИКАНСЬКІЙ ПРЕСІ

Стаття присвячена дослідженню сполучуваності лексико-семантичних груп прикметників з іменниками в англійській та американській пресі. Розглянуто залежність між частотою вживання та χ^2 . Встановлені сильні, середні та слабкі зв'язки ЛСГ прикметників з іменниками на основі статистичної обробки отриманих даних.

Як засвідчують лінгвістичні дослідження, "навіть чи в оточуючий нас дійсності знайдеться більш складне і суперечливе явище, ніж мова, в природі якої, примхливо переплітаючись, знаходять своє переломлення і відображення основні діалектичні категорії, закони об'єктивної дійсності, людського пізнання і загальноісторичного досвіду: співвідношення матеріального й ідеального, динамічна взаємодія форми і змісту, діалектичне протиріччя об'єктивного і суб'єктивного, загального й окремого, абстрактного і конкретного" [1: 36].

Саме тому на сучасному етапі розвитку семантики особлива увага приділяється вивченню не тільки власне семантики, але й синтагматичних характеристик лексичних одиниць (ЛО), які виражаються через словосполучення. Для визначення відношень між словами в словосполученнях використовують терміни валентність, "сполучуваність". Валентність – здатність слова створювати навколо себе певні вільні позиції, що займаються певними обов'язковими та факультативними актантами, тобто є лише потенційною сполучуваністю без реалізації сполучувальних задатків слова [2: 24]. Сполучуваність – його здатність вступати з іншими словами у сполучення за певними синтаксичними моделями [3: 28].

Проблемою сполучуваності слів займалися такі відомі лінгвісти, як Н.Н. Амосова, А.А. Уфимцева, Д.М. Шмельов, П. Гребе, В.В. Левицький, М.Д. Степанова, С.Д. Кацнельсон, а закони сполучуваності досліджували В.Г. Гак, М.А. Гойхман, Ю.Д. Апресян та ін. Питання сполучуваності прикметників з іменників вивчалось в роботах Л.В. Бистрової, Н.Д. Капатрука, О.Н. Бондіної, М.О. Шехтмана. Вивченням сполучуваності прикметників з іменниками в газетно-публіцистичних текстах у німецькій мові займалась В.М. Баскевич, однак сполучуваність прикметників з іменниками в англійській та американській пресі ще не була об'єктом дослідження. Тому детальніше зупинимось на особливостях сполучуваності в англійській публіцистиці.

Метою нашої наукової розвідки є дослідження сполучуваності лексико-семантичних груп (ЛСГ) прикметників з іменниками в англійській пресі за допомогою статистичних методів та впливу стильового фактору на силу сполучувального зв'язку. Ще Б. де Куртене зазначав, що потрібно застосувати в мовознавстві кількісне, математичне мислення [4: 17]. Статистичні методи вже давно широко використовуються у мовознавстві для забезпечення точності дослідження при вивченні та описі лінгвістичних об'єктів, що обумовлюється системним характером мови, якій притаманні кількісні ознаки, кількісні характеристики [5: 11]. Для використання квантитативних методів для дослідження тих чи інших об'єктів досить, "щоб ці об'єкти демонстрували повторюваність, періодичність своїх якостей, володіли тією чи іншою мірою інваріантними відношеннями, мали закономірний поділ своїх параметрів [6: 42-43].

У результаті суцільної вибірки з газетно-публіцистичних текстів англійських та американських газет (the Times, the Independent, the Guardian, the Washington Post, the New York Times, the Los Angeles Times), опублікованих у період 2001-2002 рр., було отримано 10137 синтагм, що включають словосполучення прикметників з іменниками в атрибутивній та предикативній функціях.

Робота включала наступні етапи:

1. На основі отриманої таблиці про частоту прикметників і відповідних синтагматичних зв'язків з іменниками були зроблені альтернативні (чотирипільні) таблиці для подальшої статистичної обробки [7: 102-103]. Взято до уваги сполучення, що зустрічалися більше ніж 5 разів [8: 31].

2. Зіставлення альтернативних (чотирипільних) таблиць.

3. Статистичний аналіз чотирипільних таблиць: встановлення χ^2 -критерію з метою виявлення статистично значущих зв'язків.

4. Визначення за допомогою коефіцієнта взаємної спряженості (K) ступеня взаємозв'язку між ЛО.

5. Обробка та інтерпретація отриманих даних.

Для визначення наявності чи відсутності зв'язку між ознаками використовують критерій χ^2 [9: 31]. Мінімальною величиною χ^2 , при якій зв'язок між ознаками вважається суттєвим, є величина $\chi^2 > 3,84$, при $df=1$. Випадки, коли $\chi^2 < 3,84$, вважатимемо статистично несуттєвими і в таблицю не вносимемо. Величини χ^2 і коефіцієнта спряженості K обчислювалися за формулами, які вказані у [9: 31-32]. Якщо за допомогою критерію χ^2 можна визначити, чи є суттєвим чи несуттєвим перевищення емпіричних частот над тими, що очікуються, то K показує ступінь цього зв'язку.

За отриманими результатами відзначимо найбільші показники χ^2 (з урахуванням K) для наступних семантичних підкласів прикметників з іменниками: "вік, час" з підкласом іменників "відрізки часу, доби,

періоду" – 3094,05 (K=0,552); "зовнішність людини" з підкласом "частини тіла" – 2271,367 (K=0,473); "емоційний стан та почуття" з підкласом "люди взагалі, за віковою і статевою та родинною ознаками" – 5065,727 (K=0,707); "емоційний стан та почуття" з підкласом "власні назви, імена прізвиська" – 521,312 (K=0,229); "національна приналежність" з групою іменників "явища та поняття суспільно-політичного життя" – 232,2 (K=0,151); "національна приналежність" з підкласом "збірні позначення живого" – 111,993 (K=0,105).

На перший погляд може здатися, що чим вища частота вживання, тим більшим є показник χ^2 . Однак нами відзначені випадки, коли при низькій частоті вживання є значні величини χ^2 : "зовнішність людини" з підкласом "частини тіла" – 2271, 367 (K=0,473) з частотою вживання 9; "колір та яскравість" з підкласом "одяг" – 459,932 (K=0,213) при частоті вживання 10 та з підкласом "найменування закладів, будівель, приміщень" – 1156,65 (K=0,338) з частотою вживання 82.

І, навпаки, у деяких випадках спостерігається висока частота вживання і відносно низький χ^2 : підклас прикметників "за географічним розташуванням" з підкласом іменників "географічні об'єкти, адміністративні одиниці" ($\chi^2=8,19$ K=0,028) при частоті вживання 178; підклас прикметників "за сферою діяльності об'єкта" з підкласом "дії, процеси, перетворення, факти, вчинки" ($\chi^2=9,56$ K=0,031) при частоті вживання 473; підклас прикметників "за сферою діяльності об'єкта" з підкласом "найменування закладів, будівель, приміщень" ($\chi^2 = 7,21$ K= 0,027) при частоті вживання 147.

Щоб доказово встановити, чи спостерігається залежність між частотою вживання окремих підкласів прикметників з іменниками та величиною χ^2 , ми присвоїли отриманим величинам ранги і використали для обчислень коефіцієнт рангової кореляції [9: 27]. Це дало змогу за допомогою статистичної процедури простежити, наскільки пов'язані частота вживання окремих підкласів прикметників та величина χ^2 . Розглянемо залежність на прикладі двох підкласів прикметників, один із яких "за сферою діяльності об'єкта" має найбільшу кількість корелюючих пар (8), та підклас прикметників "емоційний стан", де зафіксована найбільша величина χ^2 з кількістю корелюючих пар (4). Рангова кореляція, що була проведена з підкласом прикметників "за сферою діяльності об'єкта" та корелюючими підкласами іменників, становить від'ємну величину 0,244. Що ж стосується кореляції з підкласом "емоційний стан" з відповідними підкласами іменників, то при $df=4$ та $P=0,05$ рангова кореляція мала б дорівнювати 0,754, проте отриманий нами коефіцієнт відповідає 0,4. Результати рангової кореляції в обох випадках свідчать про відсутність прямої залежності між частотою вживання та χ^2 .

Ступінь зв'язку між досліджуваними явищами вираховується за допомогою коефіцієнта взаємної спряженості K (коефіцієнт Чупрова), між -1 до $+1$ [9: 33], який ми використали, щоб установити та розмежувати зв'язки на "слабкі", "середні" та "сильні".

Оскільки в нас сила зв'язку коливається від 0,019 до 0,7, тоді умовно приймається, що

- а) величини K = менше 0,06 відповідають слабкому зв'язку;
- б) величини від K = 0,06 до 0,15 відповідають середньому зв'язку;
- в) величини K, що більше 0,15, свідчать про наявність сильного зв'язку.

Отже, наступні ЛСГ прикметників з іменниками мають **сильний зв'язок**:

- 1.) "вартість, значення, якості та функції" з ЛСГ "абстракції": *divine inspiration, full transparency* (K=0,158);
- 2.) "вік, час" з групою іменників "відрізки часу, доби, періоду, пори року": *last month, next year* (K=0,552);
- 3.) "національна належність" з ЛСГ іменників "явища та поняття суспільно-політичного життя": *British government, the Israeli election* (K=0,151);
- 4.) "форма, обсяг, величина" з іменниками "збірні позначення живого": *a narrow majority, a tiny minority* (K=0,167);
- 5.) "форма, обсяг, величина" з ЛСГ "кількість, обсяг": *the total amounts, an average score* (K=0,18);
- 6.) "колір та яскравість" з підкласом іменників "заклади, будівлі, приміщення": *orange prison, the White House* (K=0,338);
- 7.) "колір та яскравість" із ЛСГ "одяг": *black jacket, green uniforms* (K = 0,213);
- 8.) "емоційний стан та почуття" з іменниками, що позначають людей взагалі, за віковою, статевою та родинною ознаками": *we are happy to reach an agreement, he was embarrassed and deeply disappointed* (K=0,707);
- 9.) "емоційний стан та почуття" з ЛСГ "власні назви, імена, прізвиська": *Mr.Blair has become more self-confident* (K = 0,229);
- 10.) "зовнішність людини" з групою іменників що позначають "частини тіла": *an oval face, a long hair* (K = 0,473).

Спостерігається **середній зв'язок** між наступними ЛСГ:

- 1.) ЛСГ прикметників "за сферою діяльності об'єкту" з підкласом іменників "матерія, речовина": *a nuclear bomb, chemical agents* (K=0,091);
- 2.) "національна належність" з групою іменників "збірні позначення живого": *Taliban forces, the Japanese public, German troops* (K=0,105);
- 3.) "інтенсивність, міра" з групою іменників "економічні поняття": *annual sales, double taxation* (K=0,120);

- 4.) "поведінка та дії у відношенні до об'єкта" з групою іменників "люди взагалі, за віковою, статевою та родинною ознаками": *he finds himself not only ignored* (K=0,135);
- 5.) "поведінка та дії у відношенні до об'єкта" з ЛСГ іменників "власні назви, імена, прізвиська": *Moscow remained skittish, an ungrateful Washington* (K=0,146);
- 6.) "риси характеру, звички та інші властивості людини" з групою іменників "власні назви, імена, прізвиська": *prompt Al Gore, Mr. Clinton was bald and confident* (K=0,106);
- 7.) "риси характеру, звички та інші властивості людини" з підкласом іменників "люди взагалі": *a cheerful and social boy, he is sceptical* (K=0,135);
- 8.) "зовнішність людини з ЛСГ іменників "люди взагалі, за віковою, статевою та родинною ознаками": *he has turned a sickly white, a slight pale* (K=0,102);
- 9.) "за сферою діяльності об'єкта" з ЛСГ "люди за професією, посадою, видом занять, національністю, місцем проживання" *medical specialists, biological experts* (K=0,063);
- 10.) "за сферою діяльності об'єкта" з іменниками, що позначають "назви організацій, компаній, угруповань": *a presidential committee, a militant group, a terrorist organisation* (K=0,083);
- 11.) "за сферою діяльності об'єкту" з іменниками "документи, терміни, угоди, програми": *chemical, biological and nuclear weapons programs, court documents* (K = 0,088);
- 12.) "національна належність" з ЛСГ іменників "люди за професією, посадою, видом занять, національністю, місцем проживання": *an American senator, British ambassadors, Russian doctors* (K= 0,066);
- 13.) "адміністративні, регіональні, класові розмежування з ЛСГ іменників "явища та поняття суспільно-політичного життя": *public opposition, state terrorism, provincial administration* (K = 0,066);
- 14.) "порівняння об'єкта з ознаками інших предметів" з групою іменників "дії, процеси, перетворення, факти, вчинки": *a real change, a false choice, true intentions* (K = 0,066);
- 15.) "порівняння об'єкта з ознаками інших предметів" з ЛСГ іменників "абстракції": *the true nature of price pressures, the real value, the real reason* (K = 0,076);
- 16.) "просторова оцінка щодо протяжності, відстані, тривалості" ЛСГ іменників "відрізки часу, доби, періоду, пори року": *a long time, a short term, a full year* (K = 0,065);
- 17.) "за складом, матеріалом об'єкта" з групою іменників "неживі предмети, речі, механізми": *acid bomb, paper tickets* (K = 0,061).

Можемо констатувати **слабкий зв'язок** між ЛСГ прикметників та іменників:

- 1.) "за сферою діяльності об'єкта" з підкласом іменників "дії, процеси, перетворення, факти, вчинки": *financial research, market expectations, military action* (K = 0,031);
- 2.) "за сферою діяльності об'єкта" з ЛСГ іменників "заклади, будівлі, приміщення": *chemical factories* (K = 0,027);
- 3.) "за сферою діяльності об'єкта" з ЛСГ іменників "неживі предмети, речі, механізми": *military helicopter, medical equipment* (K= 0,036);
- 4.) "за сферою діяльності об'єкта" з групою іменників "положення, стан, становище": *a financial position, an economic boom* (K=0,036);
- 5.) "вартість, значення, якості та функції" з ЛСГ "робота, праця, продукт діяльності" *important practice, base production* (K=0,038);
- 6.) "вік, час" з ЛСГ іменників "явища та поняття суспільно-політичного життя": *early elections, a new mandate* (K= 0,019);
- 7.) "національна належність" з підкласом "неживі предмети, речі, механізми": *an American plane, a French tanker* (K= 0,034);
- 8.) "форма, обсяг, величина" з ЛСГ іменників "економічні поняття": *an average income, the huge liabilities* (K= 0,03);
- 9.) "форма, обсяг, величина" з групою іменників "назви організацій, компаній, угруповань": *a global company, the broadest coalition* (K= 0,023);
- 10.) "форма, обсяг, величина" з ЛСГ іменників "заклади, будівлі, приміщення": *an average room, a circular building* (K= 0,048);
- 11.) "адміністративні, регіональні, класові розмежування" з ЛСГ іменників "люди за професією, посадою, видом занять, національністю, місцем проживання": *federal investigators, the presidential adviser, provincial leaders* (K= 0,055);
- 12.) "адміністративні, регіональні, класові розмежування" з іменниками що позначають "заклади, будівлі, приміщення": *a civilian airport, local councils* (K= 0,045);
- 13.) "порівняння об'єкта з ознаками інших предметів" з ЛСГ іменників "робота, праця, продукт діяльності": *false reports, nefarious activities* (K= 0,032);
- 14.) "за географічним розташуванням" з ЛСГ іменників "назви організацій, компаній, угруповань": *the European commission, Asian organisation* (K= 0,028);
- 15.) "за географічним розташуванням" з ЛСГ іменників "географічні об'єкти, адміністративні одиниці": (K= 0,028); *the Mediterranean island, the Atlantic coast*

- 16.) "інтенсивність, міра" з підкласом "дії, процеси, перетворення, факти, вчинки": *the sudden conflicts, intense activity, excessive charges* (K= 0,049);
- 17.) "інтенсивність, міра" з іменниками що позначають "абстракції": *zero tolerance, maximum transparency* (K= 0,047);
- 18.) "інтенсивність, міра" з ЛСГ іменників "положення, стан, становище": *increasing frustration, the gradual recovery* (K= 0,026);
- 19.) "просторова оцінка щодо протяжності, відстані, тривалості" з ЛСГ іменників "дії, процеси, перетворення, факти, вчинки": *the shortest search, endless investigations* (K= 0,019);
- 20.) "просторова оцінка щодо протяжності, відстані, тривалості" з ЛСГ іменників "економічні поняття": *long-term dividend, a continuing liability* (K=0,026);
- 21.) "колір та яскравість" з підкласом іменників "неживі предмети, речі, механізми": *black gates, white bag* (K=0,022);
- 22.) "емоційний стан" з ЛСГ іменників "люди за професією, посадою, видом занять, національністю, місцем проживання": *the inspectors were happy, shocked ministers* (K=0,052);
- 23.) "емоційний стан" з ЛСГ іменників "люди, пов'язані певними взаємовідносинами": *irritating passengers, his friends were furious* (K=0,056);
- 24.) "поведінка та дії у відношенні до об'єкта" з підкласом "дії, процеси, перетворення, факти, вчинки": *merciless hunt for terrorists, the adverse effects* (K=0,032);
- 25.) "поведінка та дії у відношенні до об'єкта" з ЛСГ іменників "люди за професією, посадою, видом занять, національністю, місцем проживання": *an implacable opponent, the predatory oligarch* (K=0,03);
- 26.) "поведінка та дії у відношенні до об'єкта" з ЛСГ іменників "назви організацій, компаній, угруповань": *companies are more reluctant, company is adamant* (K=0,038);
- 27.) "за складом, матеріалом об'єкту" з ЛСГ "робота, праця, продукт діяльності" *the metal products, steel products* (K=0,059);
- 28.) "за складом, матеріалом об'єкту" з ЛСГ іменників "матеріали, речовини": *uranium compounds, a liquid substance* (K=0,054);
- 29.) "природний та фізичний стан" з ЛСГ "заклади, будівлі, приміщення": *sealed offices* (K=0,022);
- 30.) "природний та фізичний стан" з ЛСГ іменників "географічні об'єкти, адміністративні одиниці": *stained streets, the city was tense* (K=0,025);
- 31.) "природний та фізичний стан" з ЛСГ іменників "неживі предмети, речі, механізми": *burned-out vehicles, sunken tanker* (K=0,039);
- 32.) "природний та фізичний стан" з ЛСГ іменників "положення, стан, становище": *nervous breakdown, the situation was tense* (K=0,035);
- 33.) "природний та фізичний стан" з ЛСГ іменників "люди, пов'язані певними взаємовідносинами": *freed hostages, asphyxiated hostages* (K=0,043);
- 34.) "природний та фізичний стан" з ЛСГ "люди взагалі, за віковою, статевою, родинною ознаками": *pregnant mother, hundreds of people were made homeless* (K=0,054);
- 35.) "риси характеру, звички та інші властивості людини" з ЛСГ "люди за професією, посадою, видом занять, національністю": *the suave dean, Russians are skeptical* (K=0,038);
- 36.) "риси характеру, звички та інші властивості людини" з ЛСГ іменників "збірні позначення живого": *tolerant nations, heroic defence forces* (K=0,019);
- 37.) "риси характеру, звички та інші властивості людини" з ЛСГ іменників "люди, пов'язані певними взаємовідносинами": *customers were good-humoured, the staunchest supporters* (K=0,040);
- 38.) "релігія, раса" з підкласом "люди за професією, посадою, родом занять": *Muslim scholar, Islamic extremists* (K=0,038);
- 39.) "релігія, раса" з іменниками, що позначають "абстракції": *Muslim fundamentalism, Islamic charities* (K=0,029);
- 40.) "швидкість" з ЛСГ іменників "дії, процеси, перетворення, факти, вчинки": *a slow start, a quick reaction* (K=0,047);
- 41.) "оцінка розумових здібностей людини" з групою іменників "люди за професією, посадою, родом занять": *a serious-minded pilot* (K=0,039).

Нами було встановлено 11 сильних зв'язків. Високу величину коефіцієнта K у ЛСГ "вік, час" у сполученні з ЛСГ іменників "відрізки часу, доби, періоду, пори року (K=0,552); "зовнішність людини" з "частинами тіла" (K=0,473) можна пояснити тотожністю семантичних ознак, і це цілком закономірно, тому що прикметники та іменники, які входять до вказаних лексико-семантичних груп, мають спільні сем. Сильний синтагматичний зв'язок ЛСГ "колір та яскравість" з підкласом іменників "заклади, будівлі, приміщення" (K=0,338) та з підкласом іменників "одягу" (K=0,213) обумовлюються відсутністю протилежних сем. Сполучуваність ЛСГ "емоційний стан та почуття" з підкласом іменників, що позначають людей взагалі, за віковою, статевою та родинною ознаками (K=0,707), а також з "власними назвами іменників, прізвиськами" (K=0,229) є природним, тому що людям притаманний емоційний стан, певна реакція на подію, стосунки, тощо, тобто між цими підкласами є максимальна схожість семантичних ознак. Це означає, що інтенсивність сполучуваності зумовлена виключно мовними факторами.

У наступному нашому дослідженні ми розглянемо семантичну сполучуваність ЛСГ прикметників з іменниками, проаналізуємо синтагматичні зв'язки та визначимо ступінь сполучуваності кожної ЛСГ прикметників, буде встановлено основні риси тих ЛСГ прикметників, які мають найбільший коефіцієнт вибірковості.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ТА ЛІТЕРАТУРИ

1. Уфимцева А.А. Лексическое значение: [Принцип семасиологического описания лексики].- М.: Наука, 1986. – 239 с.
2. Левицкий В.В., Огуй О.Д., Кійко Ю.С., Кійко С.В. Апроксимативні методи вивчення лексичного складу. – Чернівці: Рута, 2000. – 136 с.
3. Кочерган М.П. Слово і контекст (Лексична сполучуваність і значення слова). – Львів: Вища школа, 1980. – 184 с.
4. Бодуэн де Куртенэ И.А. Избранные труды по общему языкознанию. – Т. 1-2. – М.: Изд.-во АН СССР, 1963. – Т1. – 348 с. – Т II. – 391 с.
5. Головин Б.Н. Язык и статистика. – М.: Просвещение, 1971. – 192 с.
6. Садовський В.Н. Основания общей теории систем. – М.: Наука, 1974. – 280 с.
7. Данилович О.Д. Кількісні параметри сполучуваності прикметників з іменниками в англійській пресі // Науковий вісник Чернівецького університету. – Вип. 206-207. – Чернівці, 2004. – С.98-112.
8. Перебийніс В.І. Статистичні методи для лінгвістів. – Вінниця: Нова книга, 2002. – 172 с.
9. Левицкий В.В. Статистическое изучение лексической семантики. – К.: УМК ВО, 1989. – 155 с.
10. Бистрова Л.В. Вивчення синтагматичних зв'язків слів за допомогою статистичних методів // Мовознавство. – 1978. – № 4. – С. 44-48.
11. Гак В.Г. К проблеме семантической синтагматики // Проблемы структурной лингвистики: Сб.ст. –М.: Изд-во Моск. ун-та, 1972. – С. 367-385.

Матеріал надійшов до редакції 17.04.2006 р.

Данилович О.Д. Статистические характеристики сочетаемости лексико-грамматических групп прилагательных с существительными в английской и американской прессе.

Статья посвящена исследованию сочетаемости лексико-семантических групп прилагательных с существительными в английской и американской прессе. Рассмотрено зависимость между частотностью употребления и χ^2 . Были установлены сильные, средние и слабые связи ЛСГ прилагательных с существительными и на основании статистической обработки полученных данных.

Danylovych O.D. Statistic characteristics of valency of lexical-semantic groups of adjectives with nouns in English and American press.

The article studies the peculiarities of adjectives valency in English and American press with the help of χ^2 . The dependence between the usage frequency and the quantity of χ^2 was not confirmed. According to the intensity the force of connection between the lexical-semantic groups of adjectives with nouns by means of K were divided into the groups with high intensity, middle intensity and low intensity.