

linguistic descriptions of ancient written records, Ukrainian phonetics and dialectology problems.

*Key words:* Ukrainian linguistics, science centers, immigration, linguistic research.

УДК 811.112.2'38

Поліна ДУК

## ДОСЛІДЖЕННЯ ЛІНГВІСТИЧНИХ ОБ'ЄКТІВ МЕТОДАМИ КОРЕЛЯЦІЙНОГО АНАЛІЗУ

У статті розглянуто можливості вивчення лінгвістичних об'єктів за допомогою методів кореляційного аналізу на прикладі сучасних німецьких гумористичних віршів. Методи математичної статистики, зокрема, кореляційного аналізу дозволяють здійснити типологію лінгвістичних об'єктів з поліструктурним набором характеристик (як якісних, так і кількісних) за сукупністю всіх ознак.

**Ключові слова:** кореляційний аналіз, типологія лінгвістичних об'єктів, гумористичні вірші, текст, тип тексту, підтип тексту.

Класифікація одиниць мови за однією з їх ознак має велике поширення в лінгвістичних дослідженнях і не становить особливої складності. Так, у лінгвістиці тексту та лінгвостилістиці класифікація текстів традиційно здійснюється на підставі однієї з текстових ознак. Як приклад можна навести групу малих комічних форм: існують класифікації анекдотів за тематикою, за базовими концептами [9]; епіграми класифікують за тональністю [7], гумористичні вірші також можуть бути класифіковані за їхніми тональними характеристиками, за способом реалізації пуанта, за композиційною структурою [3] і т. п.

У той самий час усі лінгвістичні об'єкти мають набір різнорідних ознак (як якісних, так і кількісних). Традиційні для філології методи найчастіше не дозволяють зробити типологію великого масиву лінгвістичних об'єктів за сукупністю численних ознак. На нашу думку, така типологія можлива лише за застосування методів математичної статистики з використанням комп'ютерних технологій. Особливості лінгвістичної практики вимагають розробки методів кількісного опису лінгвістичних явищ,

що дозволяють точно реєструвати не стільки кількісні, скільки якісні фактори.

Розроблена нами методика класифікації лінгвістичних об'єктів, заснована на методах кореляційного аналізу з використанням ЕОМ, є одним з можливих підходів до упорядкування великих масивів фактичного матеріалу. За допомогою цієї методики виконано класифікацію німецькомовних гумористичних віршів, що враховує одночасно усі виявлені мовностилістичні характеристики аналізованих текстів.

Ця класифікація має на меті довести правомірність гіпотези Е. Різелі про можливість виокремлення найнижчого рівня ієрархічної стилістичної системи – стилю підтипів тексту, що існує в рамках стилю типу тексту [8]. У межах функціонального підходу до досліджуваного явища ми розглядаємо сукупність усіх текстів, що належать до типу тексту гумористичний вірш як складно організовану систему, що перебуває у взаємодії з аналогічними системами близьких типів тексту малих літературних форм.

Класифікація здійснюється із застосуванням деяких методів теорії нечітких множин, що припускає можливість відносити досліджувані одиниці не тільки до одного з типів, але й, у випадку перехідного характеру одиниць, одночасно до декількох типів. Така класифікація доцільна, коли границі між типами мають нечіткий, перехідний характер, що певною мірою властиво німецьким гумористичним віршам як міжжанровому утворенню. Класифікація здійснюється в кілька етапів. Насамперед, у багатьох класифікаціях постає проблема оптимального вибору системи вихідних показників, що повинна всебічно (тією мірою, якою це зумовлено суттю завдання) описувати досліджувані явища. Ці показники не повинні дублювати один одного, бути похідними один від іншого і т. д. У результаті мовностилістичного аналізу виявлені основні характеристики гумористичних віршів – їх стильові риси і варіанти їхнього прояву, покладені в основу типології. До них належать:

1) *віршованість* – поетична організація мовного матеріалу (виділені найбільш властиві види метру – ямб, трохей, анапест, а також рими – aabb, abab, abba);

2) *стислість* – відповідно до кількості рядків вірші були поділені на такі групи: 4-8 рядків, 9-12 рядків, 13-16 рядків, 17-24 рядки, 25-32 рядки, 33-40 рядків і більше 40 рядків.

3) *пуантованість* – наявність ослабленого гумористичного пуанта як результату взаємодії стилістичних прийомів гумору і засобів експресивності. Такий пуант може бути реалізований за рахунок алогізмів на значеннєвому чи на мовному рівні;

4) *тричленність* – характерна риса більшості віршів (у більш як 90% текстів виділяються три тексти, не рахуючи заголовка, – експозиція, блок хибного очікування, гумористична кінцівка); двочленну структуру мають менше 10% віршів;

5) *озаглавленість* (більшість віршів мають заголовок). Залежно від ролі у створенні гумористичного ефекту виділено 4 типи заголовків – номінативний (негумористичний) і гумористичні: проспективний, ретроспективний і проспективно-ретроспективний;

б) *гумористична тональність* вірша, яка виникає в результаті створення комічного ефекту за допомогою сукупності стилістичних засобів гумору і виразності. Виявлено близько 30-ти характерних засобів, які для здійснення цієї класифікації згруповані в такий спосіб: синтаксичні засоби, що включають фігури скорочення, фігури розчленування і фігури повтору синтаксичної структури; лексичні засоби і тропи.

Така риса, як *тривіальність*, є обов'язковою для всіх текстів, і форми її вираження не піддаються поділу на доступну для огляду кількість видів (наприклад, як це можна зробити стосовно заголовків), отже, її врахування не вплине на результати класифікації. Беручи до уваги цю обставину, ми не включаємо тривіальність у число класифікаційних ознак.

Як правило, кожна з шести стильових рис, на основі яких здійснюється типологія, проявляється у конкретному вірші в одному з декількох можливих варіантів. Наприклад, щодо такої стильової риси, як озаглавленість, кожен вірш можна віднести до однієї з 5 груп: 1 група – заголовок відсутній, 2 група – наявний номінативний заголовок, 3 група – наявний гумористичний проспективний заголовок, 4 група – заголовок ретроспективний, 5 група – заголовок проспективно-ретроспективний. Так само кожен конкретний вірш має саме такий метр, а не інший, містить певну комбінацію стилістичних засобів гумору і т. п.

Серед наведених стильових рис гумористичних віршів лише одна – стислість – вимагає особливого підходу, оскільки має кількісні параметри (кількість рядків). Щоб забезпечити єдиний підхід при здійсненні типології й мати можливість урахувати означену стильову рису, необхідно виділити групи гумористичних віршів за розмірами.

В теорії статистики існує кілька прийомів поділу статистичного ряду на інтервали [5, 1]. Для вибору оптимального принципу виділення груп віршів за кількістю рядків та їх меж нами проведено аналіз розподілу значень досліджуваного показника за допомогою гістограми. найбільш доцільним є поділ гумористичних віршів на сім груп за принципом “природних розривів”: 4-8 рядків, 9-12 рядків, 13-16 рядків, 17-24 рядки, 25-32 рядки, 33-40 рядків і більше 40 рядків.

Другим етапом класифікації є створення бази даних та розрахунок показників подібності текстів на основі виділених ознак. Мета дослідження й особливості його об'єкта – множина віршів і поліструктурний набір їхніх характеристик, тобто велика кількість і розмаїтість показників – зумовили використання прийомів кореляційного аналізу з застосуванням комп'ютерних технологій.

В основу об'єднання всього корпусу текстів у групи покладено один із принципів теорії нечітких множин, а саме принцип агрегації об'єктів на підставі максимальної гомогенності об'єктів усередині груп і максимальної гетерогенності груп між собою [4; 12]. Стосовно гумористичних віршів цей принцип припускає виділення груп на основі мінімального розходження мовностилістичних характеристик текстів, тобто найбільшої подібності за всім комплексом ознак усередині кожної з виділених груп і максимального розходження груп віршів між собою. Такий вид класифікації, при якій угруповання об'єктів здійснюється за принципом їх максимальної подібності у рамках групи та максимальної різниці між групами, прийнято називати типологією [10, с. 685].

Як міра подібності типологізованих об'єктів застосовуються різні коефіцієнти кореляції. У нашому випадку найбільш доцільним є обчислення тетрагоричного показника ( $r_{++}$ ), оскільки

класифікаційні ознаки не мають кількісного вираження, а носять якісний характер, до того ж число градацій за кожною ознакою дорівнює двом (наприклад, наявність чи відсутність заголовка певного типу, наявність чи відсутність певного метру тощо). Цей коефіцієнт кореляції є універсальним і застосовується для типології об'єктів дослідження в різних науках [11; 2]. Тетрахоричний показник обчислюється за формулою [6]:

(1)

$$r_{++} = \frac{\alpha}{\sqrt{(\alpha + \beta)(\alpha + \gamma)}}$$

де  $\alpha$  – наявність ознаки у першого і у другого вірша (+ +);  
 $\beta$  – наявність ознаки у першого вірша за відсутності її у другого (+ -);  
 $\gamma$  – відсутність ознаки у першого вірша за її наявності у другого (- +).

Для виконання автоматизованої класифікації була сформована база даних із використанням табличного процесора EXCEL. Вона містить масив даних з усіх проаналізованих текстів про наявність чи відсутність в них кожної з виділених ознак і варіантів їх прояву. За наявності певного варіанта прояву тієї чи іншої ознаки в таблицю вносилися одиниця, за відсутності – нуль.

За таким алгоритмом для кожного вірша були розраховані тетрахоричні показники  $r_{++}$  відносно всіх інших досліджуваних віршів, тобто попарно перебрано всі аналізовані тексти. Результатом стала кореляційна матриця, елементами якої власне є тетрахоричні показники  $r_{++}$ .

Наступний етап типології – власне групування віршів за результатами розрахунків. Перш за все зроблено аналіз кореляційної матриці. Якісна оцінка показників  $r_{++}$  здійснювалась на основі шкали Чеддока, яка дозволяє говорити про ступінь схожості об'єктів кореляційного аналізу за чисельними показниками [6, с. 342]. Вірші, що мають із тим чи іншим ядром показник  $r_{++}$  більше 0,7 (тобто високий і дуже високий ступінь подібності), однозначно входять до шуканої групи. Тетрахоричний

показник вірша, менший за 0,5 стосовно всіх ядер, указує на перехідний характер цього тексту. Такі тексти становили 15% від загальної кількості. Питання віднесення такого вірша до тієї чи іншої групи може бути вирішене на основі окремого порівняльного аналізу тексту. Вірші, що мають із тим чи іншим ядром показник  $r_{++}$  менший 0,3, однозначно не входять до відповідної групи.

У результаті виділено 4 підтипи німецьких гумористичних віршів. На основі тетрахоричного показника німецькі гумористичні вірші були віднесені до одного з цих 4 підтипів. У кожній групі віршів залежно від величини показника  $r_{++}$  виділяється “ядро” групи (вірші з найбільш високим показником – 0,7 і більше, тобто найтипівіші для цієї групи), “ореол ядра” (вірші з показниками 0,5 – 0,7) і “периферія групи” (вірші з низьким показником подібності – менше 0,5, можливо, що знаходяться на межі двох чи декількох груп).

Опис кожного з чотирьох підтипів німецькомовних гумористичних віршів, а також аналіз механізму реалізації пуанта в кожному з них, здійснені на заключному етапі типології, виявили різницю в прагматичних принципах конструювання пуанта, а також в реалізації інших стильових рис, що свідчить про ефективність та достовірність розробленої методики. Це відкриває нові перспективи досліджень у цьому напрямку. Типологія німецьких гумористичних віршів повинна бути розширена з урахуванням етимологічних, віршових і тональнісних параметрів творів. А запропонована типологія гумористичних віршів на основі їхнього кореляційного аналізу може бути застосована для розробки типологій інших типів текстів, а також для класифікації інших лінгвістичних об'єктів з поліструктурним набором характеристик.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Ерина А. М., Пальян З. О. Теория статистики: Практикум/ А. М. Ерина, З. О. Пальян. – К.: О-во “Знання”, КОО, 2001. – 267 с.
2. Ефимова Н. В. Практикум по общей теории статистики/ Н. В. Ефимова. – 2-е изд. – М.: Финансы и статистика, 2006. – 336 с.
3. Дук П. О. Лінгвостилістична характеристика німецьких гумористичних віршів: Автореф. дисс. ... канд. філол. наук:

10.02.04 / Харківський національний університет імені В.Н. Каразіна/ П. О. Дук. – Харків, 2007. – 20 с.

4. Заде Л.А. Основы нового подхода к анализу сложных систем и процессов принятия решений // Математика сегодня/ Л.А. Заде. – М.: Знание, 1974. – С. 5-49.

5. Лысенко С. Н., Дмитриева И. А. Общая теория статистики/ С. Н. Лысенко, И. А. Дмитриева. – М.: Форум, 2006. – 208 с.

6. Общая теория статистики: Учебник / Под ред. О. Э. Башиной, А. А. Спирина. – 5-е изд. – М.: Наука, 1999. – 384 с.

7. Перлина Ю.Г. Композиционно-стилистическая структура эпиграммы как типа текста [на материале немецких эпиграмм 17-20 вв.]: Дисс. ...канд. филол. наук/ Ю.Г. Перлина. – Днепропетровск, 1996. – 251 с.

8. Ризель Э.Г. К вопросу об иерархии стилистических систем и основных текстологических единиц // Ин.яз.в школе/ Э.Г. Ризель. – 1975. – № 6. – С. 8-15.

9. Слышкин Г.Г. Современный русский анекдот: функции и ценностные доминанты//Аксиологическая лингвистика: игровое и комическое в общении: Сб. науч. тр. / Под ред. В.И.Карасика, Г.Г.Слышкина. — Волгоград: Перемена, 2003. — С. 190 — 203.

10.Философский энциклопедический словарь / Гл. редакция: Л.Ф. Ильичев, П. Н. Федосеев, С.М. Ковалев, В.Г. Панов. – М.: Сов. Энциклопедия, 1983. – 840 с.

11.Фишер Р.А. Статистические методы для исследований/ Р.А. Фишер. – М.: Статистика, 1957. – 321с.

12.Yee Leung. Basic issues of fuzzy set theoretic spatial analysis // Pap. Reg. Sci. Assoc/ Yee Leung. – 1985. – № 58. – P. 35-46.

П. ДУК

## **ИССЛЕДОВАНИЕ ЛИНГВИСТИЧЕСКИХ ОБЪЕКТОВ МЕТОДАМИ КОРРЕЛЯЦИОННОГО АНАЛИЗА**

Статья рассматривает возможности изучения лингвистических объектов при помощи методов корреляционного анализа на примере современных юмористических немецкоязычных стихотворений. Использование методов математической статистики, в частности, корреляционного анализа позволяет создать типологию лингвистических объектов с полиструктурным набором характеристик (как качественных,

так и количественных), которая учитывает все характеристики в совокупности.

*Ключевые слова:* корреляционный анализ, типология лингвистических объектов, юмористическое стихотворение, текст, тип текста, подтип текста

P. DUK

## **THE STUDY OF LINGUISTIC OBJECTS THROUGH CORRELATION ANALYSIS METHODS**

This paper considers the possibility of applying correlation analysis methods to linguistic objects based on the example of contemporary German humorous poems. When dealing with linguistic objects that exhibit a multi-structural set of qualitative and quantitative characteristics, the use of mathematical (statistical) methods, and in particular of the correlation procedure, allows for the construction of a typology, which simultaneously takes into account all the linguistic characteristics of the objects.

*Key words:* correlation analysis, typology of linguistic objects, humorous poem, text, text type, text subtype.

УДК 821.161.2 (045)

**Людмила КОБЕЦЬ**

## **ФУНКЦІОНУВАННЯ ЛЕКСЕМИ КАМКА У ПИСЕМНИХ ПАМ'ЯТКАХ УКРАЇНСЬКОЇ МОВИ XIV – ПЕРШОЇ ПОЛОВИНИ XVII СТОЛІТТЯ**

У статті розглядається походження та лексико-семантичні особливості іменника **камка**, який зафіксований у писемних пам'ятках української мови XIV – першої половини XVII ст.

**Ключові слова:** писемні пам'ятки, українська мова, XIV–XVII ст., тематична група, запозичення.

Для дослідження історії розвитку української лексики у контексті слов'янського лексичного фонду слід приділити увагу окремим тематичним групам, які функціонували у певний історичний період. Однією з таких тематичних груп є «назви тканин».