

Федорів Я. Р.

ОСОБЛИВОСТІ ВИКОРИСТАННЯ МАТЕМАТИЧНОГО АПАРАТУ ПІД ЧАС ЛІНГВІСТИЧНИХ ДОСЛІДЖЕНЬ

Статтю присвячено висвітленню окремих проблем, які можуть виникати при некоректному застосуванні математичного апарату під час проведення лінгвістичних досліджень. Автор наводить окремі приклади, що ілюструють можливість одержання дослідником хибних висновків внаслідок поверхового трактування особливостей математичної формалізації, і пропонує шляхи їх уникнення.

Взаємодія різних галузей наук у процесі наукових досліджень є загальновідомим фактом. Прикладом такої взаємодії може слугувати взаємне проникнення і використання понять, термінів, категорій і методів з точних наук у гуманітарні, і навпаки. Зокрема, формальний апарат лінгвістики широко використовує теоретичний та дослідницький інструментарій математики, математичної логіки, статистики та кібернетики, як-от: *одиниця, множина, модель, мережа, поле, фрейм, класифікація, алгоритм, функція, значення, відношення, обмеження, предикатно-аргументна структура, проблема багатозначності, методи розв'язання багатозначності, оптимізація, інформація, кодування/декодування* тощо. Можна також навести приклади наслідків зворотного термінологічного руху, наприклад, з лінгвістики в кібернетику: *моделювання мови, штучна мова, синтаксична будова мови, метамова, машинний переклад, розпізнавання і синтез мовлення, текстові та словникові бази даних, діалогові системи* тощо.

З цього погляду уявляється цілком природним факт використання математичного апарату від час виконання теоретичних та експериментальних досліджень у галузі гуманітарних наук **взагалі** й лінгвістики зокрема. Більше того, у **наш** час застосування у лінгвістичних дослідженнях математичного апарату вважається не--мінною запорукою точності й достовірності дослідження, чого бракує звичайним, «вербальним» міркуванням і доказам. При цьому іноді, не розбираючись як слід у математичному апараті, деякі дослідники схильні приписувати йому надприродну могутність, особливо коли необхідні розрахунки зроблено за допомогою комп'ютера. Після класичної фрази: «Дані заклали в кібернетичну машину, і вона...» впливають висновки, які нібито заслуговують на тим більшу довіру, чим складніші були перетворення

і чим більше машинного часу витрачено. Методи оголошуються «точними» тільки тому, що в них використано математичний апарат.

З такою точкою зору не можна погодитись. Математичний апарат сам по собі не надає дослідженню точності, а його висновкам — доказовості. Істинна точність у гуманітарних науках не пов'язана з математичною формалізацією, а досягається глибиною проникнення в матеріал, повнотою врахування історичних, літературних, лінгвістичних та інших зв'язків. Некоректно використаний математичний апарат у лінгвістиці відіграє роль чогось на зразок бойового розфарбування дикуна: надаючи видимої «точності» формальному дослідженню в неформальній галузі, він, по суті, мало чим допомагає, а іноді навіть виявляється програвшим.

Для potwierдження цієї думки можна навести приклади деяких хибних висновків, зроблених у галузі лінгвістики внаслідок поверхового ставлення до використовуваного математичного апарату. Так, не кожен лінгвіст-дослідник знає, що за допомогою статистичних методів не можна довести, що такий-от текст належить такому-то автору, наприклад, Шевченку; максимум, що можна довести, це те, що наявний матеріал не суперечить гіпотезі щодо авторства Шевченка. Іншими словами, статистичні методи дозволяють довести, що певна гіпотеза не абсурдна, що не слід її одразу відкидати, проте довести, що вона справедлива, не можна.

Однак у процесі розв'язання проблеми встановлення авторства тексту можна простежити, як відбуваються різні маніпуляції з матеріалом, підраховуються і порівнюються характеристики ритміки текстів, будуються криві розподілу довжини слів (у складах і буквах) і робляться відповідні висновки. Якщо у порівнюваних текстах за всіма параметрами спостерігаються розбіжності, то слід враховувати, що ці розбіжності можуть

бути як випадковими, так і значущими. Тому для оцінки значущості дослідники застосовують відомий у математичній статистиці критерій Колмогорова—Смирнова. Проте необхідно пам'ятати, що будь-який математичний критерій слід застосовувати коректно. Так, критерій Колмогорова—Смирнова може застосовуватися тільки до неперервних випадкових величин; у даному ж випадку ці величини (число складів, число букв у слові) дискретні, тобто їхні можливі значення (1, 2, 3...) розділені проміжками. Крім того, умовою застосування будь-якого критерію узгодженості є наявність визначеного числа дослідів (спостережень) у статистичному матеріалі, причому передбачається, що ці досліді незалежні. Нарешті, досліджуваний матеріал повинен бути однорідним, інакше до нього взагалі не можна застосовувати імовірнісні методи.

Ya. R. Fedoriv

PECULIARITIES OF USING MATHEMATICAL TOOLS IN LINGUISTIC INVESTIGATIONS

The paper discusses some problems caused by inadequate ways of using mathematical tools in linguistic investigations. The author analyses some examples of superficial treatment of mathematical formalization peculiarities, leading to fallacious inferences, and suggests the ways of avoiding erroneous results.

Взагалі, точність і результативність будь-якого математичного дослідження в прикладній галузі залежить від того, як побудована математична модель досліджуваного явища, чи адекватна вона реальності, як тлумачаться результати розрахунків, зроблених за цією моделлю. Іноді ряд очевидних питань, що стосуються певної математичної моделі, не тільки не розв'язується, але навіть не ставиться.

Таким чином, дослідник, обираючи математичний апарат, насамперед повинен розуміти, що саме він робить, і викласти це так, аби зрозумів читач. Застосовувана математична модель повинна бути чітко описана, прийняті допущення — сформульовані таким чином, щоб не виникло сумнівів щодо якості та коректності як власне математичної частини роботи, так і всього дослідження.