

## ВІДГУК ОФІЦІЙНОГО ОПОНЕНТА

кандидата технічних наук

**Остапової Ірини Вадимівни**

на дисертаційну роботу **Хомицької Ірини Юріївни**

«Методи та засоби диференціації фоностатистичних структур  
функціональних стилів англійської мови»

поданої на здобуття наукового ступеня кандидата технічних наук  
за спеціальністю 10.02.21 – Структурна, прикладна та математична  
лінгвістика

*Актуальність теми.* Збільшення текстового матеріалу зумовлене об'єктивними чинниками, з-поміж яких найголонішим є використання інформаційних технологій в усіх сферах людської діяльності. В умовах глобалізації щоденно опрацьовується великий обсяг текстової інформації. Зрозуміло, що із збільшенням кількості текстів в інтернеті, зростає відсоток текстів, авторство яких потрібно встановити. Найчастіше встановлення авторства тексту стосується сфери судочинства (анонімні листи, заповіти та інші судові документи); сфери науки (встановлення відсотку авторського внеску статей, що мають більше ніж одного автора); офіційно-ділової сфери (документи) та сфери літератури (художні твори). Незважаючи на певні напрацювання у цьому напрямі, завдання не є розв'язаним повною мірою. Підвищення достовірності встановлення авторства тексту є актуальним, бо методи, які застосовуються для розв'язання даного завдання не дають вичерпної інформації для повної характеристики авторського стилю. Авторка дисертаційної роботи пропонує підвищити достовірність встановлення істотних відмінностей між авторськими стилями, використовуючи поєднання статистичних критеріїв. У цьому випадку отриманий результат

підтверджується кількома статистичними критеріями, що дає змогу констатувати істотні відмінності між текстами різних авторів з більшою достовірністю. Тому з погляду сучасного стану досліджень, розроблення методу комплексного аналізу, що ґрунтується на використанні декількох статистичних критеріїв є актуальним науковим завданням. Отже, тема дисертаційного дослідження “Методи та засоби диференціації фоностатистичних структур функціональних стилів англійської мови” є актуальною.

*Зв’язок роботи з науковими програмами, планами, темами.* Основні результати, що представлені в роботі, були отримані в рамках виконання держбюджетних науково-дослідних робіт Національного університету “Львівська політехніка”, а саме: тема “Застосування формальних моделей для дослідження інформаційних природномовних процесів” (номер державної реєстрації 0114U001239); “Нейромережева технологія захисту та передачі даних у реальному часі з використанням шумоподібних кодів” (номер державної реєстрації 0119U002256), 2019-2020 рр.

*Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертаційній роботі.*

Наукові положення, результати і висновки дисертаційної роботи є загалом достатньо обґрунтованими. Обґрунтованість та достовірність отриманих наукових результатів забезпечено чіткістю і послідовністю постановки завдань дослідження та коректним застосуванням теоретично обґрунтованих класичних методів. Автором запропоновані нові методи і вдосконалені існуючі моделі та засоби для диференціації фоностатистичних структур стилів англійської мови. Висновки чітко висвітлюють одержаний ефект від проведених досліджень, що підтверджено результатами апробацій на 8 наукових конференціях.

*Наукова новизна отриманих результатів.* До основних наукових результатів, одержаних здобувачкою особисто належать:

1) вперше розроблено метод комплексного аналізу диференціації фоно-статистичних структур стилів англійської мови, який ґрунтується на поєднанні статистичних методів гіпотез, ранжування і визначення відстаней між стилями та дає змогу підвищити достовірність диференціації стилів;

2) вперше розроблено багатофакторний метод визначення ступеню дії факторів стилю, підстилю та авторської манери викладу, який ґрунтується на зіставленні фоностатистичних структур різних авторів, різних підстилів і різних стилів та забезпечує можливість підвищення достовірності авторської атрибуції тексту за його загальною стилевою маркованістю;

3) вдосконалено статистичну модель стилевої, підстилевої та авторської диференціації за методом гіпотез та методом ранжування, яка відрізняється від існуючих розмежуванням текстів за кількістю груп приголосних фонем, що уможливорює підвищення достовірності визначення диференційних ознак стилів, підстилів та текстів різних авторів;

4) вдосконалено статистичну модель визначення стилерозрізняльної здатності груп приголосних фонем, яка на відміну від інших, дає змогу визначити групу приголосних фонем з найвищою стилерозрізняльною здатністю і, тим самим, зменшити кількість груп приголосних фонем, за якими розмежовуються стилі, що підвищує достовірність та рівень автоматизації диференціації текстів.

*Практичні результати роботи, їх рівень та ступінь впровадження.* Основні результати дисертаційного дослідження застосовано для методичного забезпечення дисциплін “Основи прикладної лінгвістики” і “Контрастивна стилістика англійської та української мов” у навчальному процесі кафедри

прикладної лінгвістики Національного університету “Львівська політехніка”, що засвідчено відповідним актом.

Отримані результати мають практичну цінність, яка полягає у наступному: розроблено структуру системи диференціації фоностатистичних структур стилів англійської мови, яка ґрунтується на модульному принципі, що дає змогу швидко модифікувати і вдосконалювати програмну систему; розроблено алгоритми реалізації методів та моделей диференціації фоностатистичних структур стилів англійської мови для ефективнішого розмежування стилів на фонологічному рівні; розроблено програмні засоби системи автоматизованої диференціації фоностатистичних структур стилів на мові програмування Java, що забезпечує кросплатформність програмного засобу та дає змогу автоматизувати процес авторської атрибуції тексту.

*Повнота викладення результатів досліджень в опублікованих працях.*

За результатами дисертаційного дослідження опубліковано у 23 наукових працях, з яких: 8 статей у наукових фахових виданнях України, 5 з яких включені до міжнародної наукометричної бази Index Copernicus, 2 статті у науковому періодичному виданні України, що входить до міжнародної наукометричної бази Scopus, 1 стаття у науковому періодичному виданні Швейцарії, що входить до міжнародної наукометричної бази Scopus, та 12 публікацій у матеріалах конференцій, 11 з яких входять до міжнародних наукометричних баз Scopus та Web of Science.

Публікації достатньою мірою відображають основні результати роботи та зміст дисертації.

*Оцінка основного змісту дисертації та її структури.* Дисертаційна робота є науково-дослідною роботою, яка характеризується завершеністю, чітким та легким для сприймання стилем викладу матеріалу дослідження. Основними складовими дисертації є анотація, вступ, чотири розділи,

висновки, список використаної літератури та додатки.

У *вступі* окреслено основний напрям проведення дослідження, відповідно до якого обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету та основні завдання для її досягнення, розкрито наукову новизну роботи, зазначено практичне значення отриманих результатів. А також розглянуто зв'язок роботи з науковими темами і особистий внесок авторки, наведено інформацію про апробацію результатів роботи та публікації.

У *першому розділі* проведено аналіз існуючих методів, моделей та програмних засобів диференціації стилів та авторської атрибуції текстів.

На думку здобувачки, у дослідженнях, в яких стилі диференціюються на фонологічному рівні, рівень достовірності не є достатнім, бо істотні відмінності між текстами не підтверджуються результатами отриманими кількома статистичними критеріями. Тому розроблення методу комплексного аналізу, який ґрунтується на поєднанні статистичних критеріїв, є актуальним. Підвищення достовірності авторської атрибуції забезпечується розробленим багатофакторним методом визначення ступеню дії факторів стилю, підстилю та авторської манери викладу, який є відсутній у сучасних дослідженнях.

Дисертантка зазначає, що для підвищення достовірності стилевої диференціації, необхідно вдосконалити існуючі статистичні моделі, застосовуючи потужний непараметричний критерій Колмогорова-Смірнова та диференціювати стилі за кількістю груп приголосних фонем. Для підвищення рівня достовірності та автоматизації стилевої, підстилевої та авторської атрибуції тексту, необхідно вдосконалити статистичну модель визначення стилерозрізняльної здатності груп приголосних фонем досліджуваних стилів на основі визначення групи фонем з найбільшою стилерозрізняльною здатністю та диференціювати стилі за цією групою фонем.

У плані аналізу засобів диференціації текстів здобувачка зазначає, що існуючі програми опрацьовують тексти на лексико-семантичному рівні, структура якого є менш строгою ніж фонологічного рівня (транскрипційні символи), і його важче формалізувати. Тому достовірність результатів на лексико-семантичному рівні є нижчою, ніж на фонологічному рівні. Зважаючи на це, актуальним є розроблення програмної системи авторської, підстилевої і стилевої атрибуції тексту на фонологічному рівні.

У другому розділі дисертантка описує розроблені методи диференціації функціональних стилів англійської мови, зокрема: *метод комплексного аналізу диференціації фоностатистичних структур стилів та багатофакторний метод визначення ступеню дії факторів стилю, підстилю та авторської манери викладу.*

Для підвищення достовірності розмежування стилів, підстилів та текстів різних авторів, розроблено метод комплексного аналізу диференціації фоностатистичних структур стилів, який поєднує такі статистичні методи: метод гіпотез (критерії Стюдента, Колмогорова-Смірнова та хі-квадрат), метод ранжування та метод визначення відстаней між стилями. Підвищення достовірності диференціації стилів забезпечується застосуванням п'яти методів, одним з яких є непараметричний і його потужність прямує до 1 (критерій Колмогорова-Смірнова). Результати отримано з довірчою ймовірністю 0,95.

Для підвищення достовірності встановлення стилевої, підстилевої та авторської приналежності досліджуваного тексту, розроблено багатофакторний метод визначення ступеню дії факторів стилю, підстилю та авторської манери викладу, який передбачає визначення загальної стилевої маркованості досліджуваного тексту (за запропонованою формулою).

Особливістю методу є розмежування текстів за значенням загальної стилевої маркованості зіставлюваних текстів. Розроблений метод реалізується за схемою стиль→підстиль→автор.

*У третьому розділі представлено вдосконалені статистичні моделі: модель стилевої, підстилевої та авторської диференціації текстів за методом гіпотез і ранжування та модель визначення стилерозрізняльної здатності груп приголосних фонем досліджуваних стилів.*

Для розв'язання завдання підвищення достовірності визначення диференційних ознак стилів, підстилів та текстів різних авторів, вдосконалено статистичну модель стилевої, підстилевої та авторської диференціації текстів за методами гіпотез і ранжування, яка, на відміну від існуючих, враховує позицію фонем у слові, визначає відношення кількості груп приголосних фонем, за якими встановлено істотні відмінності між попарно зіставленими стилями до загальної кількості груп приголосних фонем. Ідея розмежування стилів за кількістю груп приголосних фонем пропонується вперше. Велика кількість груп приголосних фонем (6 – 8), за якими тексти відрізняються істотно, дає змогу з більшою достовірністю робити висновки про їх відмінність.

Для розв'язання завдання підвищення рівня автоматизації диференціації стилів, вдосконалено статистичну модель визначення стилерозрізняльної здатності груп приголосних фонем, яка базується на визначенні відношення кількості порівнянь, за якими виявлено істотні відмінності між стилями – до кількості всіх порівнянь. Зменшення кількості груп приголосних фонем, за якими диференціюються стилі, спрощує процес атрибуції тексту, підвищує його достовірність і забезпечує вищий рівень автоматизації шляхом здійснення атрибуції тексту за групою приголосних фонем з найбільшою стилерозрізняльною здатністю.

У четвертому розділі описано розроблену програмну систему автоматизації диференціації фоностатистичних структур функціональних стилів англійської мови та наведено результати дослідження.

Розроблено структуру програмної системи диференціації фоностатистичних структур функціональних стилів англійської мови, яка ґрунтується на модульному принципі, що дає змогу швидко модифікувати розроблене програмне забезпечення (ПЗ), підвищити рівень автоматизації та ефективність стилевої, підстилевої та авторської атрибуції тексту. Розроблене ПЗ є кросплатформним, оскільки розроблення системи здійснено на об'єктно-орієнтованій мові програмування Java. Побудована структура класів ПЗ системи диференціації стилів.

*Висновки* відображають досягнуту мету дослідження та повною мірою висвітлюють отримані в роботі результати.

*Список використаної літератури* є інформативним, достатньо повно охоплює предметну область та відображає опрацювання автором значної кількості іноземних джерел.

***Відповідність дисертації та автореферату встановленим вимогам.*** За своєю структурою, обсягом і оформленням дисертація та автореферат відповідають вимогам, встановленим до кандидатських дисертацій.

Автореферат за змістом ідентичний основним положенням, що викладені в дисертації, та не містить інформації, яка не відображена в самій роботі.

Стиль викладу матеріалів досліджень, наукових положень і рекомендацій забезпечує їх адекватне і належне сприйняття.

***Загальна оцінка роботи.*** Дисертація та автореферат написані грамотно, послідовно, коректно та мають завершену логічну структуру. Поставлену автором мету досягнуто, сформульовані завдання розв'язано, висновки повністю відображають основний зміст роботи.



*До недоліків та зауважень дисертаційної роботи можна віднести:*

1. В дисертації та авторефераті зустрічається термін “автоматизація”, а значення рівня автоматизації, який забезпечує розроблений програмний продукт - відсутнє.
2. Здобувачка розробила моделі та методи, але відсутня інформація про параметри цих моделей та методів (економічність).
3. На мою думку, необхідно дати додаткові пояснення щодо вибору груп приголосних фонем за критерій розмежування стилів (ст. 45) та пояснити формулювання “неврахована позиція фонем у слові” (ст. 101).
4. Окрім того, слід зазначити та додатково пояснити відмінність між формулюваннями “тексти різних авторів” і “різні авторські стилі” (ст. 134).
5. В дисертації, відсутня інформація про можливість розмежування стилів за однією і тією самою групою фонем у всіх зіставленнях (ст. 80).
6. Здобувачу в дисертації, на мою думку, варто додатково обґрунтувати вибір текстового матеріалу наукового стилю (ст. 126).
7. Відсутня інформація про можливість розмежування стилів в одній і тій самій позиції фонем у слові у всіх зіставленнях (ст. 76).
8. Відсутній опис та призначення бази даних текстів на рис.4 в авторефераті та дисертації.
9. На мою думку, варто було дати короткий опис перспективи застосування описаних методів до текстів українською мовою.

Однак зазначені зауваження не знижують загальної позитивної оцінки дисертаційної роботи.

### *Висновки*

В цілому, дисертаційна робота Хомицької І.Ю. на тему “Методи та засоби диференціації фоностатистичних структур функціональних стилів англійської мови” є завершеною науковою працею, яка розв’язує актуальне наукове завдання розроблення і вдосконалення методів, моделей і засобів диференціації фоностатистичних структур функціональних стилів англійської мови та відповідає спеціальності 10.02.21 – Структурна, прикладна та математична лінгвістика.

Автореферат відповідає змісту дисертації та оформлений згідно з вимогами.

Дисертаційне дослідження є актуальним, містить наукову новизну і практичну цінність та відповідає вимогам до дисертаційних робіт, а саме: пп. 9, 11, 12 «Порядку присудження наукових ступенів». Тому вважаю, що його авторка Хомицька Ірина Юріївна заслуговує присудження їй наукового ступеня кандидата технічних наук за спеціальністю 10.02.21 – Структурна, прикладна та математична лінгвістика.

Офіційний опонент

Старший науковий співробітник

Українського мовно-інформаційного фонду

НАН України,

кандидат технічних наук



І.В. Остапова

Власноручний підпис
Засвідчую:
Завідувач канцелярії НАН України
«13» квітня 2017 р.