

Довгун І. М.

Національний університет харчових технологій, Київ, Україна

Статистичні характеристики голосних і приголосних морфонеми англійської мови в текстах трьох функціональних стилів

Для дослідження особливостей функціонування голосних і приголосних морфонеми англійської мови в мовленні було вирішено взяти тексти трьох функціональних стилів: художнього, наукового й газетного. Вибір стилів зумовлений своєрідністю вживання лексичних і граматичних одиниць у них. Так, художній стиль використовує весь потенціал мови, він багатий на засоби вираження, у ньому представлені слова з емоційним забарвленням, слова, вжиті в переносному значенні, порівняльні звороти. Метою художнього тексту є вплив на людську психіку, емоції, почуття. Науковий стиль передбачає точне повідомлення зі спеціальної галузі науки, в ньому часто вживаються терміни, запозичення з інших мов. Газета – найсвіжіше джерело інформації, у ній присутні риси як художнього, так і наукового стилів. Зазначені стилі відібрані для порівняння, наскільки активно морфонеми функціонують у них.

Функціональні властивості будь-якої мовної одиниці визначаються особливостями її вживання у певному тексті, при цьому великого значення набувають правила організації вибірки та визначення обсягу матеріалу з метою одержання статистично вірогідних результатів. При формуванні вибірки важливо враховувати її репрезентативність, тобто вона повинна рівномірно розподілятися по всій генеральній сукупності та мати досить великий обсяг. Усі вибрані тексти написані англійською мовою і є хронологічно обмеженими.

Для аналізу кожного стилю було відібрано по 20 вибірок обсягом 250 лексичних одиниць методом зональної вибірки. Щоб рівномірно представити досліджувані тексти у вибірках зонами в нас були їх початок, середина та кінець. Кожне слово, в якому представлена морфонема, записувалося на окрему картку, що містила інформацію про слово, його морфемне членування й тип морфонеми.

Для визначення статистичних характеристик голосних і приголосних морфоне́м англійської мови в текстах трьох функціональних стилів зробимо ряд обчислень: знайдемо середню частоту морфоне́м в кожному масиві й межі її коливання.

Середня частота голосних морфоне́м у художньому стилі дорівнює 47,55, у газетному – 29,2, у науковому – 13, тобто середня частота морфоне́м коливається від 13 до 47,55 [див. табл. 1]. Чи є суттєвими ці розходження та які межі коливання частоти? Для того, щоб дати відповідь на це запитання, встановимо такі статистичні характеристики: середнє квадратичне відхилення (σ), міру коливання середньої частоти ($\sigma_{\bar{x}}$).

Таблиця 1

Статистичні характеристики голосних морфоне́м на 5000 словоформ

Стиль	\bar{x}	σ	$\sigma_{\bar{x}}$	$\bar{x} - 2\sigma_{\bar{x}}$	$\bar{x} + 2\sigma_{\bar{x}}$	$S \frac{2}{x}$
Художній	47,55	10,7	2,39	42,77	52,33	6,03
Газетний	29,2	10,04	2,25	24,7	33,7	5,31
Науковий	13	4,01	0,9	11,2	14,8	0,85

Використовуючи вираховані статистичні характеристики голосних морфоне́м, з'ясуємо, чи залежить їх функціонування від функціонального стилю, тобто перевіримо на нульову гіпотезу. Ми формулюємо таку нульову гіпотезу: розходження між частотами у порівнюваних вибірках не істотне, воно обумовлене статистичними законами, а не характером досліджуваних текстів. Якщо наші підрахунки покажуть, що розходження не є значимими, то зможемо підтвердити нульову гіпотезу та віднести три масиви до однієї генеральної сукупності.

Спочатку зіставимо вибірки за смугою коливання середньої частоти. Для цього нанесемо на діаграму смуги коливання середньої частоти голосних морфоне́м у трьох функціональних стилях [див. рис. 1].

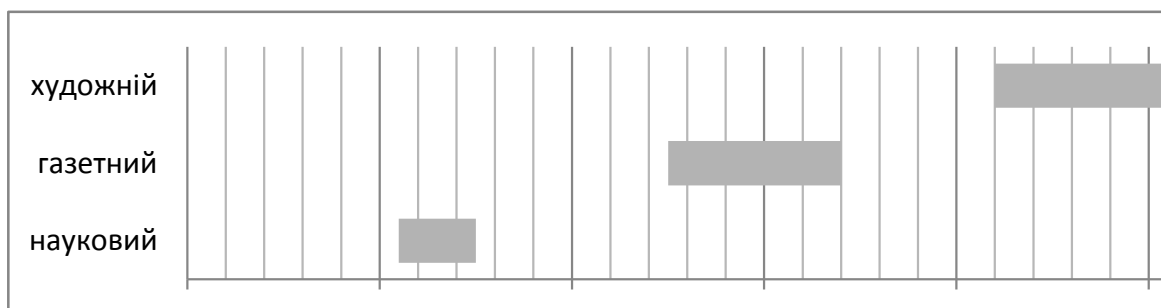


Рис. 1 Коливання середньої частоти голосних морфонем у досліджуваних текстах

На діаграмі видно, що всі три смуги не перетинаються, тому з ймовірністю 95% можна стверджувати, що всі три вибірки істотно розходяться між собою за середніми частотами функціонування голосних морфонем у текстах. Це свідчить про те, що ці тексти розмежовуються. Перевіримо отримані дані за критерієм Стюдента, тобто встановимо, чи суттєві розходження між масивами за параметром “середня частота”. Якщо коефіцієнт критерію Стюдента менший 2,04, то розходження між нашими вибірками несуттєві, якщо він вищий 3,65 – суттєві. Узагальнені дані за критерієм Стюдента наведено в Таблиці 2.

Таблиця 2

Значення критерію Стюдента для пар досліджуваних текстів

Стиль	Газетний	Науковий	Художній
Художній	5,45	-	-
Газетний	-	6,53	-
Науковий	-	-	13,17

В таблиці 2 бачимо, що показники в трьох парах масивів набагато більші критичних показників t , тому з імовірністю 99,9% ми можемо стверджувати, що розходження у частоті голосних морфонем у досліджуваних текстах істотне й не зумовлене статистичними причинами. Отже, частота голосних морфонем може бути статистичним параметром для трьох досліджуваних текстів.

Використовуючи дані за критерієм Стюдента, можемо визначити ступінь розходження, відстань (1) між зіставляваними масивами за даною ознакою. Варто зауважити, що відстань між стилями можливо вирахувати лише для тих

пар текстів, де показник критерію Стюдента свідчить про суттєве розходження між ними. Відстань між стилями вираховується за такою формулою:

$$l = \frac{t - t_a}{t},$$
 де l – відстань між зіставленими масивами текстів, t –

одержаний нами показник критерію Стюдента, t_a – критичний показник критерію Стюдента в таблиці [1, 156]. Значення l знаходиться в межах від 0,0 до 1,0. Ми приймаємо такі пороги щодо визначення відстані між досліджуваними масивами текстів:

1. Велика відстань (1,0 – 0,80);
2. Відстань вище середньої (0,79 – 0,60);
3. Середня відстань (0,59 – 0,40);
4. Мала відстань (0,39 – 0,20);
5. Мінімальна відстань (0,19 – 0,0).

Чим вищий показник l , тим менша близькість між текстами. Якщо відстань між ними велика чи вища середньої, то вони розглядаються як такі, що не належать до однієї сукупності й характеризуються різними особливостями за даним параметром. Мала й мінімальна відстань свідчать про схожість зіставлених масивів за досліджуваним параметром. Середня відстань є перехідною зоною між двома протилежностями.

Пара масивів художній-науковий знаходяться в зоні відстані вище середньої (0,72), газетний-науковий – у зоні середньої відстані (0,44), художній-газетний – у зоні малої відстані (0,33) з довірчою ймовірністю 99,9%. Отже, за параметром “середня частота” художній і науковий масиви розрізняються за середньою частотою голосних морфем. Газетний і науковий тексти знаходяться у перехідній зоні, тому важко сказати, чи досліджуване явище характеризується стилерозрізною потужністю.

Тепер виявимо спільні й відмінні риси, які характеризують досліджувані тексти за частотою приголосних морфем.

Середня частота приголосних морфонема у художньому масиві дорівнює 19,45, у газетному – 13, у науковому – 6,7, отже, їх середня частота коливається від 6,7 до 19,45 [див. табл. 3].

Таблиця 3

**Статистичні характеристики приголосних морфонема на 5000
словоформ**

Стиль	\bar{x}	σ	$\sigma_{\bar{x}}$	$\bar{x} - 2\sigma_{\bar{x}}$	$\bar{x} + 2\sigma_{\bar{x}}$	$S \frac{2}{x}$
Художній	19,45	5,43	1,21	17,03	21,87	1,55
Газетний	13	6,07	1,36	10,28	15,72	1,94
Науковий	6,7	3	0,67	5,36	8,04	0,47

Нанесемо на діаграму коливання середньої частоти приголосних морфонема у трьох масивах [див. рис. 2].

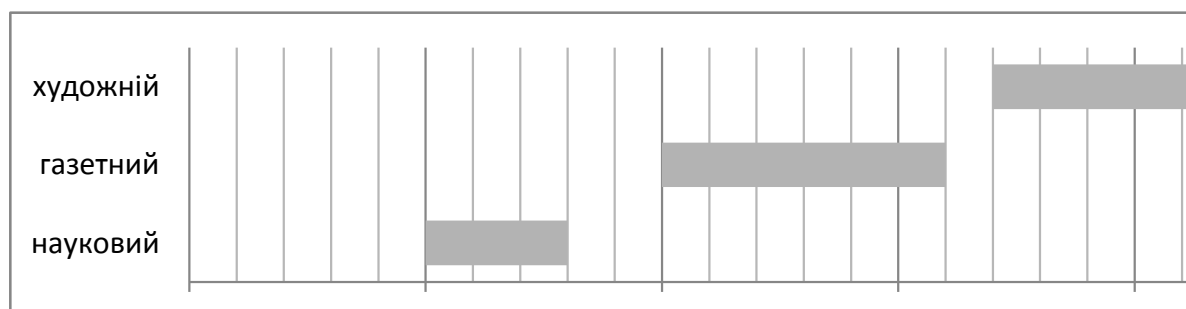


Рис. 2 Коливання середньої частоти приголосних морфонема у досліджуваних текстах

Як видно з діаграми на рис. 2, за параметром “середня частота” всі три смуги не перетинаються. Отже, досліджувані масиви розмежовуються за середньою частотою функціонування приголосних морфонема у них.

За критерієм Стюдента встановлюємо, чи суттєві розходження між даними текстами за параметром “середня частота”. Одержані дані наводяться в таблиці 4. Порівняння даних із критичним показником [1, 156] показує, що з довірчим рівнем 99,9% суттєвими є розходження між парами стилів газетний-науковий, художній-науковий, тоді як розходження між художнім і газетним масивами з рівнем 98%. Варто зауважити, що при зменшенні довірчого рівня до 99% розходження між парою художній-газетний зростає.

Значення критерію Стьюдента для пар досліджуваних текстів

Стиль	Газетний	Науковий	Художній
Художній	3,45	-	-
Газетний	-	4,06	-
Науковий	-	-	8,97

Використовуючи дані за критерієм Стьюдента, визначимо відстань (I) між зіставлюваними масивами текстів. З довірчим рівнем 99,9% для пари стилів художній-науковий показник відстані між стилями дорівнює 0,69 (відстань вища середньої), для газетного-наукового – 0,32 (мала відстань), для художнього-газетного – 0,20 (мала відстань). За прийнятими нами порогоми можемо зробити висновок, що художній і науковий масиви розрізняються за середньою частотою приголосних морфонема, так само як у випадку з голосними.

Отже, порівняння масивів показало, що середня частота голосних і приголосних морфонема може бути стилерозрізнявальною ознакою.

Література

1. Перебийніс В. І. Статистичні методи для лінгвістів : навч. посібник / В. І. Перебийніс. – Вінниця : Нова Книга, 2002. – 172 с.