

ЗАСТОВУВАННЯ МЕТОДІВ КОРПУСНОЇ ЛІНГВІСТИКИ ДЛЯ ДОСЛІДЖЕННЯ ОСОБЛИВОСТЕЙ ВИКОРИСТАННЯ СУЧАСНОГО МОЛОДІЖНОГО СЛЕНГУ

Ніфтілін В.В., Борисова Н.В.

*Національний технічний університет
«Харківський політехнічний інститут»,
м. Харків*

Сучасні досягнення корпусної лінгвістики – відносно нової, проте перспективної галузі, що з'явилася у зв'язку з інтенсивним розвитком комп'ютерних технологій, – дозволяють розширити існуючі та провести раніше неможливі дослідження різноманітних соціолінгвістичних явищ. У межах даного дослідження при вивченні такого явища як молодіжний сленг методи корпусної лінгвістики використовувались для визначення частот вживання сленгових слів та виразів, отримання інформації щодо вікових особливостей їх використання. Спочатку було створено корпус повідомлень соціальних мереж та блогів, що містили сленгові слова, а потім розроблено програму автоматизованої обробки цього корпусу. Джерелом текстів для корпусу були повідомлення у пабліках «JustStory», «Убивчі Історії» та «Веселі історії з життя людей» різних соціальних мереж. Для корпусу відбиралися не всі тексти пабліків, а тільки такі, що містили інформацію щодо віку особи, яка опублікувала пост. Корпус складається з трьох підкорпусів, у кожному з яких по 30 текстових файлів. Кожний файл містить один пост. Кількість слів у файлі коливається від 150 до 300. Розмір файлів складає 2-3 Кб. Таким чином, взагалі у корпусі міститься 15000 слів. Крім корпусу повідомлень соціальних мереж та блогів, що містять сленгові слова, для роботи програми автоматизованої обробки корпусу використовувався словник молодіжного сленгу. Обсяг словника складає близько 3000 слів. Словник було отримано у результаті анкетування користувачів соціальних мереж, що проводилося з грудня 2016 по червень 2017 року включно. В результаті роботи програми автоматизованої обробки корпусу отримуємо список з найчастіше вживаних сленгових слів та виразів, а також вік людей, які їх використовують. Сортування здійснюється за зменшенням частоти використання сленгового слова незалежно від віку особи, яка опублікувала пост. Подальший аналіз отриманої інформації дозволив сформулювати списки сленгових слів та виразів, що використовуються особами певного віку (у нашому випадку від 17 до 24 років). Наприклад, у користувачів сімнадцятирічного, двадцятирічного, двадцятидворічного та двадцятичотирьохрічного віку найпопулярнішим є слово «го» (розпочинати щось), що зустрілося відповідно 637, 521, 384, 469 разів; у вісімнадцятирічних – «анон» (анонім), 87 повторювань; у дев'ятнадцятирічних – «дру» (друїд, персонаж деяких ігор), 91 використання; у двадцятидворічних – «нік» (реєстраційне ім'я в грі, на сайті й т.п.), 198 вживань; у двадцятичотирьохрічних – «ап» (досягнення чергового рівня у грі), зустрілося 115 разів. Перспективою даного дослідження може бути вивчення часових та регіональних особливостей використання сленгових слів та виразів молоддю.