

Посты в Twitter и Reddit – снимки эволюции языка



BOOTYFUL, cyw, scirms. Знаете вы или нет, что обозначают эти слова, но в ближайшее время вы можете с ними столкнуться. Эти слова — из числа самых быстро распространяемых слов в Соединенном Королевстве, о чем говорит программное обеспечение, созданное для составления графика развития языка онлайн.

“Bootyful” - альтернативное написание слова “beautiful” (рус. красивый) - используют поразительно большое количество пользователей Twitter в Южном Уэльсе. “Cyw” - “coming your way” (рус. станет доступным для тебя; идет к тебе) - стало популярно на севере Великобритании. “Scirms” - слово, обозначающее тренировки перед соревнованиями, появилось на игровых форумах.

Программное обеспечение, которое находит подобные слова, было создано Дэниелом Кершоу и Мэттью Роу в Ланкастерском университете в Великобритании. Для создания графиков популярности слов они

использовали научно признанные методы, придуманные лексикографами. Ученые перевели эти методы в алгоритмы и применили эти алгоритмы к 22 миллионам слов в Twitter и Reddit.

Цель данного исследования — выявить путь классического английского из романов в мейнстрим и показать область интересов, в которой изменяется язык в сети. “Если мы встречаем новое слово в Twitter или Reddit, то первый вопрос, который у нас возникает — когда же оно появится в газетах?” - говорит Роу.

***“За несколько месяцев фраза “on fleek”
перешла из онлайн-видео в текст
популярной песни”***

Алгоритмы выявляют не просто популярные слова, а слова, быстро набирающие популярность, поэтому результаты исследования требуют некоторых пояснений. Например, в работе, которую планируют представить в следующем месяце на конференции The Web Search and Data Mining (рус. Веб-поиск и добыча данных), едва ли покажут полную картину эволюции языка в городах.

“Именно сейчас происходит что-то невероятное”, - говорит Кэтрин Мартин, глава US Dictionaries в Oxford University Press в Нью-Йорке. “Очень многое из того, что было бы употреблено только в устной речи и не записано, в настоящее время становится текстом постов, а впоследствии люди начинают искать эти слова в интернете и находить”.

Кэтрин Мартин говорит, что ее команда использует подобные алгоритмы поиска, чтобы находиться в курсе процесса эволюции языка в сети. “Мы просматриваем сайты, анализируя сотни миллионов английских слов, употребляемых пользователями каждый месяц в самых различных источниках. Так мы можем проследить развитие языка месяца за месяцем”.

Согласно К. Мартин, эволюцию языка ускоряет далеко не онлайн-коммуникация, коммуникация лишь позволяет легко увидеть эту эволюцию. “Язык постоянно меняется, но многие изменения происходят незаметно. Примером изменения могут быть слова, используемые при общении с детьми на севере Иллинойса, которые никто и не подумал записать”.

Интернет может также представлять собой канал для распространения новых слов по всему миру, что раньше могло полноценно выполнить только телевизионное вещание. Например, фраза “on fleek”, которая обозначает “looking good” (рус. тот, который выглядит хорошо, красиво, приятно). Впервые это выражение появилось в июне 2014 в видео в приложении Vine, опубликованного девушкой по имени Пичес Монро (Peaches Monroee). Пичес

употребила эту фразу по отношению к своим бровям. Спустя несколько месяцев фраза появилась в песне Nicki Minaj. А видео Пичес набрало около 38 миллионов просмотров.

“Водоворот новых слов, их придумывания и распространение в сети — это настоящий вызов для лингвистов”, - говорит Дэвид Бернхарт, лексикограф из Массачусетса, чьи разработки использовали в своем ПО Кершоу и Роу. Интернет отнес на второй план исследования языка романов, которые преобладали раньше, в эру печатного слова. “Сегодняшняя проблема - как же отделить по-настоящему ценные слова от бессмысленных”, - заявляет Бернхарт.

Автор статьи - Хэл Ходсон (Hal Hodson)

New Scientist №3057, опубликовано 23 января 2016

Источник:

Tweets and Reddit posts give snapshot of our changing language, By Hal Hodson// New Scientist, 20 January 2016.

URL: <https://www.newscientist.com/article/2073691-tweets-and-reddit-posts-give-snapshot-of-our-changing-language/> (дата обращения: 04.02.16)