

ПЛЕНАРНОЕ ЗАСЕДАНИЕ

СТАТИСТИЧЕСКАЯ РЕЛЕВАНТНОСТЬ КОЛЛОКАЦИЙ С КЛЮЧЕВЫМ СЛОВОМ BLACK

Д.В.Дроздова (Минск, МГЛУ)

На примере исследования синтагматических связей существительного *black* 'черный' демонстрируется, как инструменты корпус-менеджера Sketch Engine позволяют выявить статистически релевантные коллокации. Результатом исследования являются коллокации, которые отражают сложившиеся социальные стереотипы этнической группы афро-американцев в веб-дискурсе.

Ключевые слова: корпусная лингвистика, лексико-синтаксическая сочетаемость, коллокация, статистический анализ, этно-маркированные слова.

STATISTICAL RELEVANCE OF COLLOCATIONS WITH THE KEYWORD BLACK

D.V.Drozdova (Minsk, MSLU)

On the example of studying the syntagmatic relations of the noun *black*, the author demonstrates how Sketch Engine's corpus manager tools is used to identify statistically relevant collocations. The result of the study is collocations that reflect the prevailing social stereotypes of the Afro-American ethnic group in web discourse.

Key words: corpus linguistics, lexical-syntactic compatibility, collocation, statistical analysis, ethnically marked words.

Цель работы – выявить статистически релевантные коллокации с этно-маркированным словом *black* 'черный'. В фокусе нашего внимания находятся сочетания, которые обладают такими свойствами, как повторяемость, специфичность значения, структурная устойчивость и экспрессивность. На основе анализа сочетаний слов с точки зрения регулярности и частотности употребления в определенном типе текстов можно установить коммуникативно релевантные компоненты значения и проследить семантические трансформации, которые слова претерпевают в процессе употребления.

В концепциях Ш. Балли и В. В. Виноградова подобные сочетания определяются как синтагмы. Данный термин подчеркивает синтаксическую природу синтагмы как речевой единицы и ее системный характер (синтагматические отношения – одни из основных в лексической системе языка). По мнению А. А. Реформатского, «синтагмой является сочетание двух членов, связанных тем или иным отношением с неравноправной направленностью членов, где один член является определяемым, а другой – определяющим» [2, с.137].

В 1930-х годах Дж. Р. Ферт ввел термин «коллокация» для характерных или «устоявшихся» словосочетаний. Сам термин и общая идея, заключались в том, что словосочетания «соответствуют некоторому традиционному способу высказывания вещей» [1, с.15]. Данное видение коллокации было охотно подхвачено исследователями в различных областях, что привело к серьезной терминологической путанице, которая окружает понятие сегодня.

В рамках корпусной лингвистики, принят термин коллокация, который может быть определен как статистически устойчивое словосочетание [3, с.138]. Исходя из определений, приведённых выше, синтагма и коллокация обладают одинаковыми свойствами: устойчивость, наличие главного и зависимого слов (ключевого слова и коллоканта), повторяемость, в связи с чем в настоящей работе данные термины являются взаимозаменяемыми.

Актуальность темы коллокаций определяется ростом интереса лингвистов к проведению исследований на материале корпусов и перспективам применения современных статистических методов для анализа языковых явлений в новом формате.

Для выявления статистической релевантности с ключевым словом *black* 'черный' был избран подкорпус американских новостных статей корпус английского языка English Web 2021 (enTenTen21). Инструментом анализа выступил корпус-менеджер Sketch Engine [6]. Выбор данного инструментария обусловлен наличием функции вывода списков коллокаций по отдельным синтаксическим моделям с указанием силы связи между лексемами.

Программное обеспечение корпусного менеджера позволяет исследовать семантические свойства слов путем выявления контекста использования слова, что в дальнейшем может быть отражено в численном виде. По умолчанию коллокации в корпусном менеджере Sketch Engine сортируются так, чтобы наиболее типичные словосочетания находились вверху. Статистическая мера $\log\text{Dice}$ используется для определения того, насколько типичной (или сильной) является коллокация.

Исследуемое слово *black* 'черный' было использовано в качестве существительного, что сделает его ключевым для поиска коллокаций.

В числе частотных глаголов, используемых с ключевым словом *black* 'черный' преобладают глаголы действия, совершаемые над представителем афро-американской этно-расовой группы: *blacks were lynched* 'негров линчевали' (6.2 балла), *enslaved blacks* 'порабощали черных' (7.4 балла), *hate blacks* 'ненавидеть черных' (5.2 балла), *oppressed blacks* 'угнетали черных' (6.0 балла). Следует отметить наличие пассивных конструкций или пассивные значения, которые свидетельствуют о навязывании позиции слабого.

При рассмотрении прилагательных, используемых с ключевым словом *black* 'черный', было выявлено такие прилагательные как, *inferior* 'низший' (8.7 балла) указывающий на низкий социальный статус, *African* 'африканский' (7.5 баллы), *American* 'американский' (6.7 балла) указывающие на этническую принадлежность, *lazy* 'ленивый' (6.7 балла) демонстрирующий сложившийся стереотип о представителях афро-американской этно-расовой группы.

Статистические меры, используемые в корпусном менеджере, для выявления релевантных коллокаций, способствуют уменьшению затраченного времени на поиск и отбор контекстов, а также позволяет исследовать контексты в большом объеме, что делает результаты более репрезентативными. На небольшом фрагменте исследования можно увидеть важные статистические данные

о социокультурном компоненте значения слова. Таким образом, в ходе данного исследования была выявлена целесообразность и статистическая релевантность использования коллокаций с ключевым словом *black* при анализе социокультурных трансформаций (например, дискриминацию, стереотипизацию, этноцентризм, политкорректность).

СПИСОК ЛИТЕРАТУРЫ

1. Evert, S. *The Statistics of Word Cooccurrences: Word Pairs and Collocations* / S. Evert. – University of Stuttgart, 2004. – 353 p.
2. Реформатский, А. А. *Введение в языковедение* / под ред. В. А. Виноградова. – М.: Аспект Пресс, 1996 (1967). – 536 с.
3. Захаров, В. П. Анализ эффективности статистических методов выявления коллокаций в текстах на русском языке / В. П. Захаров, М.В. Хохлова // *Компьютерная лингвистика и интеллектуальные технологии*. Вып. 9 (16). – М.: РГГУ, 2010. – 138 с.
4. Виноградов В.В. *Лексикология и лексикография: Избранные труды* – М.: Наука, 1977. – 319 с.
5. Killgarriff, A. *The Sketch Engine: Ten Years On, Lexicography* [Electronic resource]: Mode of access: https://www.sketchengine.eu/wpcontent/uploads/The_Sketch_Engine_2014.pdf. – Date of access: 18.02.2024.
6. *Sketch Engine* [Electronic resource]: Mode of access: <https://www.sketchengine.eu/> – Date of access: 18.02.2024.