

КОРРЕЛЯЦИЯ МЕЖДУ ДРЕВНОСТЬЮ И УПОТРЕБИТЕЛЬНОСТЬЮ ЗНАЧЕНИЙ ДВУЗНАЧНЫХ АНГЛИЙСКИХ СЛОВ

И. А. Терентьева

Воронежский государственный университет

Поступила в редакцию 22 октября 2010 г.

Аннотация: в статье исследуется зависимость функциональной нагрузки от возраста слова.

Ключевые слова: полисемия, функциональная нагрузка многозначных слов, возраст слова, частотность слов.

Abstract: the article explores the dependence functional load of polysemantic words on the age of a word.

Key words: polysemy, functional load of polysemantic words, age of word, word's frequency.

Лексическая многозначность сама по себе достаточно хорошо изучена [1; 2; 3]. Корреляция между многозначностью и частотностью слов также попадала в поле зрения лингвистов [4; 5; 6; 7; 8; 9; 10], но связь частотности значений слова с их возрастом не была предметом специального исследования. Данная работа посвящена исследованию взаимосвязи между древностью и употребительностью значений английских слов. Объектом исследования являются частотно-семантические словари английского языка [11; 12]. Предметом исследования являются закономерности, существующие между возрастом слова и употребительностью его значений.

Рассмотрим теоретически процесс вытеснения одного значения другим. Вначале в языке было однозначное слово, и на него приходилось 100 % функциональной нагрузки. Затем появля+тся новое (вторичное) значение, и его функциональная нагрузка невелика. Если употребительность этого значения начинает расти, то автоматически уменьшается употребительность первичного значения. В итоге этот процесс может закончиться полным вытеснением первичного значения (см. рис. 1).

Значения слов на шкале времени распределены согласно данным Oxford English Dictionary [13], а относительная частота употребительности значения слова взята из словаря И. Лорджа и Э. Торндайка и указана в промилле.

На первом этапе работы мы разделили все многозначные слова на группы по количеству значений. В данной статье речь пойдет только о группе двузначных слов, так как она имеет наибольшее количество слов – 4367.

Следующий этап работы – указание для каждого значения даты его первой фиксации согласно Oxford English Dictionary.

В результате сопоставления хронологии первой фиксации значений с их относительной частотой мы получили 2112 слов, у которых функциональная нагрузка (ФН) более древнего значения выше, чем функциональная нагрузка нового значения.

У 1234 слов большую функциональную нагрузку имеют более «молодые» значения. У 918 слов значения распределяются 50/50. 84 слова имеют одинаковую дату фиксации для обоих значений. И наконец, у 17 слов значения распределяются 50/50, и они зафиксированы в один и тот же год.

Теперь рассмотрим каждую группу подробнее. В первой группе у всех слов большую функциональную нагрузку имеют более древние значения. В среднем разница между функциональной нагрузкой первого и второго значений составляет 551 ‰. Разница в возрасте между первым и вторым значениями составляет в среднем 187 лет. Максимальная разница в возрасте равна 1079 годам, а минимальная – 1 год. Мы понимаем, что эти сведения являются относительными, и дата первой фиксации не обязательно является датой первого употребления значения слова.

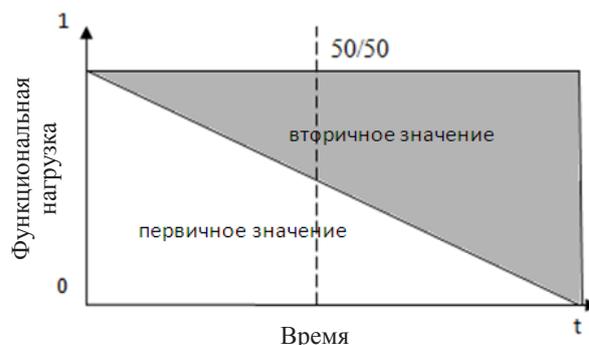


Рис. 1. Хронология изменения функциональной нагрузки на значения двузначного слова

Всего было посчитано 41 151 употребление значений двузначных слов. Для удобства анализа мы разделили все слова на 10 групп, результат показан на рис. 2.

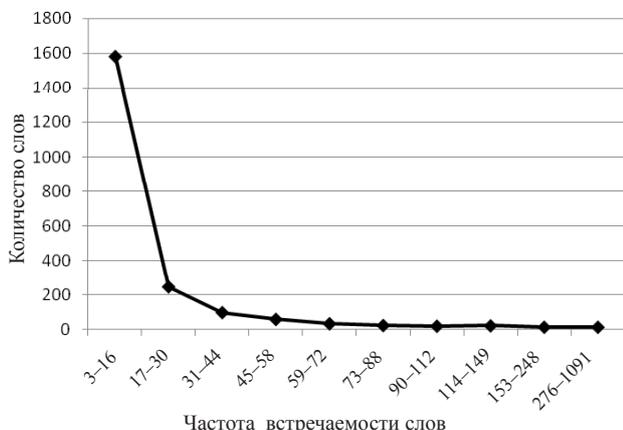


Рис. 2. Распределение употребительности слов

Как мы видим, наиболее представительная первая группа, в которую входят слова с частотой от трех до шестнадцати. (Слова с частотой 2 в ней не рассматриваются, так как они не имеют степеней свободы). В эту группу попало более половины слов от всех имеющихся, правда, их статическая достоверность невелика.

Теперь рассмотрим, как соотносится функциональная нагрузка и возраст слова. Это соотношение показано на рис. 3.

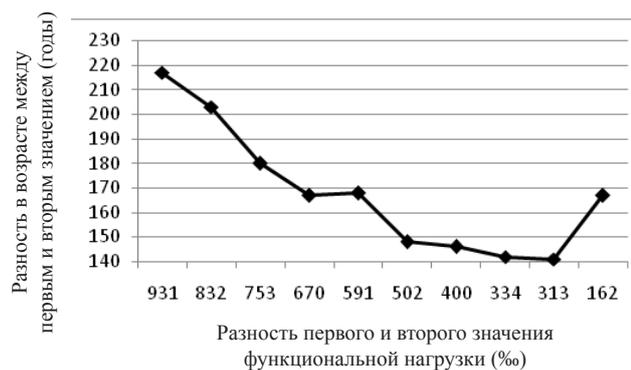


Рис. 3. Корреляция между расхождением функциональной нагрузки и возраста значений слова (для всего материала)

В целом, мы видим, что чем больше употребительность значения слова, тем более древним оно является. Неоднородность графика связана с тем, что для одних слов время фиксации совпадает со временем процесса и, следовательно, отражает последовательность появления значений, а у других слов время

фиксации удалено и не отражает последовательности процессов.

Для того чтобы получить статистически достоверные сведения, мы решили исключить из этой группы слова, которые употреблялись менее 10 раз. Таким образом, у нас осталось 918 слов. Теперь снова рассмотрим, как соотносится функциональная нагрузка и возраст слова. Это соотношение показано на рис. 4.



Рис. 4. Корреляция между расхождением функциональной нагрузки и возраста значений слова (для статистически достоверного материала)

Так же, как и в прошлый раз, общая тенденция не изменилась: чем больше расхождение в ФН между значениями, тем больше расхождение во времени первых фиксаций этих значений.

Теперь рассмотрим слова, у которых функциональная нагрузка «молодых» значений больше. Всего таких слов оказалось 1234. В процентном отношении процесс вытеснения второго значения первым составляет 28 %. В среднем разница между употребительностью первого и второго значений составляет 492 %. Разница в возрасте между первым и вторым значениями составляет в среднем 133 года. Максимальная разница в возрасте равна 1110 гг., а минимальная – 1 год.

Если обратиться к абсолютным числам, то всего было посчитано 18 955 употреблений значений двузначных слов. Мы разделили все 18 955 слов на 10 групп, результат изображен на рис. 5.

Как мы видим, снова большинство слов попадает в первую группу. Это слова, которые употребляются в одном из двух значений от трех до двенадцати раз. Максимальная встречаемость одного и того же слова равна 344, тогда как максимальная встречаемость у группы, которая подтверждает первую часть нашей гипотезы, – 1091.

После того, как мы убрали из этой группы слова, которые употреблялись менее 10 раз, остается всего 462 слова. Затем мы построили график соотношения функциональной нагрузки и возраста слов.



Рис. 5. Распределение употребительности слов



Рис. 6. Соотношение функциональной нагрузки и возраста слова

Как видно на рис. 6, всего лишь две точки находятся вдалеке от остальных, а в целом мы не наблюдаем особой динамики.

Теперь рассмотрим слова, у которых функциональная нагрузка распределяется 50/50. Всего таких слов оказалось 918. Разница в возрасте между первым и вторым значениями составляет в среднем 137 лет.

Было посчитано 2454 употреблений значений двузначных слов; как они распределяются, можно увидеть на рис. 7.

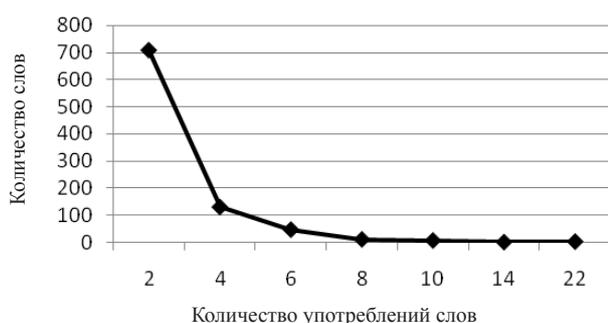


Рис. 7. Распределение употребительности слов

Большинство значений слов встретилось всего лишь два раза.

К сожалению, этих сведений недостаточно, чтобы точно сказать о функциональной нагрузке слов. Поэтому мы не будем учитывать эти слова. Таким образом, у нас остается 111 слов, оба значения которых имеют одинаковую функциональную нагрузку. Это достаточно редкий случай.

Из приведенных выше результатов следует, что наиболее многочисленной является группа слов, в которой функциональная нагрузка более древнего значения больше, в эту группу входит 48 % от общего количества слов. На втором месте находится группа с большей функциональной нагрузкой более «молодых» значений, и в эту группу входит 28 % слов. Из этого можно сделать вывод, что распределение функциональной нагрузки связано с возрастом слова: вначале значение слова с большей функциональной нагрузкой является более древним, чем значение слова с меньшей функциональной нагрузкой. Позже употребительность значений многозначного слова изменяется, и функциональная нагрузка более древнего значения может сравняться с функциональной нагрузкой более «молодого» значения. В итоге, позднее значение может приобретать большую функциональную нагрузку, чем раннее значение. В дальнейшем мы планируем проверить нашу гипотезу на материале других многозначных слов.

ЛИТЕРАТУРА

1. Андреевская А. В. Квантитативное исследование полисемии корневых слов русского языка XI–XX веков / А. В. Андреевская // Учен. зап. Тартуского ун-та. – 1990. – Вып. 912 : Квантитативная лингвистика и автоматический анализ текстов. – С. 3–11.
2. Апресян Ю. Д. Лексическая семантика. Синонимические средства языка / Ю. Д. Апресян. – М., 1974.
3. Зализняк Анна А. Многозначность в языке и способы ее представления / Анна А. Зализняк. – М., 2006.
4. Кретов А. А. Принципы выделения ядра лексико-семантической системы / А. А. Кретов // Семантика слова и синтаксической конструкции. – Воронеж, 1987. – С. 84–93.
5. Поликарпов А. А. Полисемия : системно-квантитативные аспекты / А. А. Поликарпов // Учен. зап. Тартуск. гос. ун-та. – Тарту, 1987. – Вып. 774 : Квантитативная лингвистика и автоматический анализ текстов. – С. 135–168.
6. Поликарпов А. А. Циклические процессы в становлении лексической системы языка : моделирование и эксперимент : автореф. дис. ... д-ра филол. наук / А. А. Поликарпов. – М., 1998. – 55 с.
7. Тулдава Ю. Проблемы и методы квантитативно-системного исследования лексики / Ю. Тулдава. – Таллин : Валгус, 1987. – 204 с.

8. Филиппов А. К. Распределение частот значений полисемичных глаголов (на материале ЛСГ глаголов мышления в русском языке) / А. К. Филиппов // Материалы XXXIX Международной филологической конференции. Секция прикладной и математической лингвистики. – СПб., 2010. – С. 66–73.

9. Zipf G. K. The Psycho-Biology of Language an introduction to dynamic philology / G. K. Zipf. – Cambridge : Mass. MIT Press, 1965. – 336 p.

10. Zipf G. K. Human Behavior and the Principle of Least Effort / G. K. Zipf. – Cambridge : Addison-Wesley Press, 1949. – 573 p.

11. Lorge I. The semantic count of 570 Commonest English Words / I. Lorge. – New York, 1949.

12. Lorge I. A Semantic Count of English Words / I. Lorge, E. L. Thorndike. – Columbia : Teachers College, Columbia University, 1938. — 1177 p.

13. Oxford English Dictionary on CD-ROM. Version 3.1. Oxford : Oxford University Press, 2009.

Воронежский государственный университет

Терентьева И. А., аспирант кафедры теоретической и прикладной лингвистики

E-mail: irina1985_2004@mail.ru

Тел.: 8-910-344-70-26

Voronezh State University

Terentyeva I. A., Post-graduate Student of the Theoretical and Applied Linguistics Department

E-mail: irina1985_2004@mail.ru

Tel.: 8-910-344-70-26