

Интернационализация русской лексики: количественный аспект

заимствование, интернационализм, частотность, сочетаемость

В статье описывается конкуренция заимствованной и исконной лексики в русском языке на протяжении XVIII–XX вв. При помощи лингвографических и корпусных методов проведено сравнение заимствований с их русскими эквивалентами с точки зрения семантики и лексической сочетаемости.

В этой работе при помощи современных лексикографических источников [Ляшевская 2009; Егорова 2014], созданных на основе корпусных данных, мы приведем наиболее распространённые заимствования и сравним их частотность с частотностью их исконных эквивалентов в диахронии. При этом нами будет проведён анализ различия их сочетаемости на основе Ngram-анализа сверхобъёмных корпуса текстов «Гугл Книги» [Lieberman 2007].

В результате лингвографической и корпусной работы выявлены языки-посредники для 120 наиболее частотных в русском языке заимствований-интернационализмов греческого и латинского происхождения. Частеречная разметка выявила следующий морфологический состав списка: существительные – 70 слов, глаголы – 28, прилагательные и наречия – 22.

Исследованные пары иноязычных когнат и их аналогов славянского происхождения в 37 случаях демонстрируют существенную разницу в объёме значений (рекомендация // совет) и различную сочетаемость (английская аббревиатура // резкое сокращение).

Проанализировав 120 пар слов с точки зрения частотности и сочетаемости, мы пришли к выводу о том, что лексемы славянского происхождения в разы превосходят по частотности своих конкурентов и имеют больше значений, более широкую сочетаемость.

В 13 случаях пары демонстрируют стилистические и прагматические (по сфере функционирования) различия. Заимствования чаще используют в научном и официально-деловом стилях речи, а в художественных и религиозных текстах – славянизмы (аналог теоремы // подобие Божие), но в целом исконная лексика более нейтральна по стилю: аналогичный результат (книжн.) // подобный случай (нейтр.).

Анализ сочетаемости лексем показал наличие следующих тенденций: сингармонизм в плане происхождения (например, заимствованное

существительное обычно сочетается с заимствованным же прилагательным, а глагол славянского происхождения требует аналогичного по происхождению наречия); славянизмы сочетаются с более экспрессивной лексикой и лексикой, характеризующейся аксиологией, заимствования же более нейтральны; тенденции в плане синтаксической сочетаемости с именами собственными (заимствованные глаголы часто управляют именами собственными, в то время как их исконные аналоги – нарицательными).

Расхожее мнение о мнимом засилии иностранных слов в русской лексике не подтверждается количественными характеристиками относительной частотности пар заимствование/славянизм (например, пара аргументы – 96239 / доводы – 5123712, где доля иноязычного заимствования в суммарной частоте двух лемм – всего лишь 1,84%).

При анализе влияния заимствованного слова на частотность его конкурента славянского происхождения было проанализировано соположение кривых на графиках частотности конкурирующих слов и выяснено, что в 67 случаях рост частотности интернационализма на частотность исконной леммы не влияет, а в 58 случаях можно говорить о конкуренции лексем: в 19 случаях частотность славянизма снизилась одновременно с ростом частотности заимствованного слова, ещё в 19 случаях заимствование сравнялось по частоте со славянизмом (преимущественно среди существительных), и в 20 случаях частотность заимствования превысила слово славянского происхождения (высока доля прилагательных, глаголов). Последние 2 явления (выравнивание частотности; количественное доминирование заимствований) произошли опять же на фоне снижения частотности исконной леммы.

Частотность заимствованной леммы на исход её конкуренции со словом славянского происхождения влияет следующим образом: чем выше частотность (от 10 млн словоупотреблений леммы за 200 лет), тем меньше его доля – не более 10% от суммарной частоты вариантов пары; чем реже слово встречается в русском языке (до 50 тыс. словоупотреблений леммы за 200 лет), тем выше доля заимствований в паре – 60-90% в паре.

Для каждого рассмотренного заимствования также определялось характерное время роста частоты встречаемости слова в корпусе. Данная величина показывает, как быстро заимствование вытесняет слова-аналоги. Наиболее типичное значение – медианное – составило 42 года. Для половины рассмотренных примеров характерное время роста лежит в интервале 26-98 лет. Именно столько лет нужно иностранному слову, чтобы превзойти по частотности аналог славянского происхождения в книжной речи.

Благодарности

Работа поддержана РФФИ, грант № 17-29-09163.

Литература

Ляшевская О.Н. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка) / О.Н.Ляшевская, С.А.Шаров. – М.: Азбуковник, 2009.

Егорова Т.В. Словарь иностранных слов современного русского языка / Т.В.Егорова. – М.: Аделант, 2014. – 800 с.

Lieberman E. Quantifying the evolutionary dynamics of language / E.Lieberman, J.-B.Michel, J.Jackson, T.Tang, M.A.Nowak // Nature. – 2007. – Vol. 449. – Pp. 723–716. <http://www.nature.com/nature/journal/v449/n7163/abs/nature06137.html> [Электронный ресурс. Дата обращения: 15.10.2020].