Семантические трансформации политической лексики

А.Ц. Масевич[1], В.П. Захаров [2]

[1] Санкт-Петербургский государственный институт культуры

[2] Санкт-Петербургский государственный университет

andmasev@mail.ru v.zakharov@spbu.ru

Аннотация

Представлены результаты исследования, которое проводилось на базе русского корпуса Google Books. Использовались программные инструменты Google Books N-gram Viewer, позволяющий строить графические модели частотного поведения лексических единиц — слов и коллокаций в заданный период времени, а также корпусная система Sketch Engine, с помощью которой был создан и исследовался корпус из текстов газет «Правда» и «Известия» за период с 1930 по 1939 гг.
С помощью этого корпуса, а также графиков, построенных в системе Google Books N-gram Viewer, удалось проследить, как меняются во времени коннотации некоторых лексических единиц. Описаны три примера таких изменений, показана их корреляция с историческим процессом.

Ключевые слова: корпусы текстов, диахронические исследования, частотность лексических единиц, модели частотного поведения, семантическое поле.

1. Введение

В настоящей публикации представлены некоторые результаты длительного исследования частотного поведения политической лексики в определенный временной период.

Исследование проводилось на основе русского корпуса Google Books (590 тыс. текстов русских книг, более 67 млрд. словоупотреблений). В качестве инструмента использовалась система Google books Ngram Viewer [1], детально описанная нами в нескольких публикациях [2, 3]. Система свободно доступна в Интернете с 2010 г. Она позволяет строить графики встречаемости слов и коллокаций за выбранный временной период.

На горизонтальной оси графика показываются годы, входящие в заданный временной период. По вертикальной оси откладывается выраженная в процентах относительная частота встречаемости в корпусе заданной N-граммы в соответствующем году. Относительная частота встречаемости N-граммы за определенный год подсчитывается следующим образом: число употреблений N-граммы в данном году делится на общее число словоупотреблений в корпусе в этом же году, результат умножается на 100 [1].

При построении графика имеется возможность задать ряд условий, позволяющих делать графики более наглядными, сопоставлять кривые поведения нескольких N-грамм, выявлять наиболее частотные словосочетания с данной словоформой и их поведение во времени и др.

В ходе нашего исследования построено около 500 графиков частотного поведения слов и коллокаций, относящихся к политической лексике. Часть результатов этой работы опубликованы в [4]. Кроме того полностью с результатами исследования (свыше 500 графиков с комментариями) можно ознакомиться на ресурсе Research Gate [5] [[6].

Часть исследования проводилась на базе корпусов газет «Правда» и «Известия» за 1930-е гг., работающих под управлением корпусной системы Sketch Engine [7].

В данной публикации сделана попытка показать семантические трансформации некоторых единиц политической лексики в период с 1920 по 2000 год. Под семантической трансформацией мы будем понимать в первую очередь изменение коннотаций некоторых лексических единиц.

Текст печатного документа советского времени, в особенности такой, который описывает или комментирует текущие политические ситуации, имеет все черты тоталитарного дискурса [8]. Одной из таких черт является частое употребление «застывших» формул — регулярно воспроизводимых в тексте речевых штампов. В тоталитарном дискурсе их употребление — это соблюдение речевого ритуала. Употребление «застывших» выражений становится обязательным условием речевой деятельности, осуществляемой публично, и сигнализирует о том, что такая речевая деятельность удовлетворяет «правилам игры».

2. Частотное поведение и коннотации словоформы «сталинские»

Очевидно, что изменения коннотации лексических единиц вызваны двумя причинами: во-первых, увеличением числа публикаций об определенных событиях и исторических фигурах прошлого, а во вторых, изменением отношения к этим событиям и фигурам.

Наиболее показательной исторической фигурой является И.В. Сталин, и лексика, которая связана с его правлением, представляет большой интерес для разных аспектов гуманитарного знания.

Мы провели исследование по корпусам газет «Правда» и «Известия» за 1930-е гг.[1] Корпуса созданы под управлением корпусной системы Sketch Engine

[1] Создатель корпуса — Кириллова А.О., бакалавр образовательной программы

(http://the.sketchengime.co.uk). Система позволяет строить конкорданс, дистрибутивный тезаурус, создать частотные списки разных единиц, вычислять степень устойчивости сочетаний и мн. др. Здесь мы приведем лишь некоторые данные, касающиеся употребления имен собственных и демонстрирующие частоту употребления лексических единиц с корнями «-сталин-» и «-ленин-». Нас они заинтересовали с точки зрения отражения в открытой печати существовавшего уже тогда культа вождя.

Таблица 1. Суммарная частота встречаемости словоформ с корнем «-*сталин*-» в корпусе газет «Известия» и «Правда» за 1930-е гг.

Газета	Период	Абсолютная частота	IPM
Известия	1930-1934	45	257,47
Известия	1935-1939	465	5680,64
Правда	1930-1934	72	772,39
Правда	1935-1939	198	2976,68

Таблица 2. Суммарная частота встречаемости словоформ с корнем «-*ленин*-» в корпусе газет «Известия» и «Правда» за 1930-е гг.

Газета	Период	Абсолютная частота	IPM
Известия	1930-1934	123	703,75
Известия	1935-1939	342	4178,02
Правда	1930-1934	155	1662,79
Правда	1935-1939	133	1999,49

Как видно из табл. 1, 2 наибольшее число слов с корнями «-*сталин*-» и «-*ленин*-» встречается в публикациях газеты «Известия» во вторую половину 1930-х гг. При этом следует отметить, что отмечается динамичный рост абсолютной частоты слов с корнем «-*сталин*-» во вторую половину 1930-х гг. по сравнению с первой половиной того же десятилетия, а именно: в 10,3 раза по газете «Известия» и в 2,75 раза по газете «Правда». По-другому меняется абсолютная частота слов с корнем «-*ленин*-». По газете «Известия» она возрастает в 2,8 раза, а по газете «Правда» снижается на 22 словоупотребления.

Обращают на себя внимание чрезвычайно высокие значения ipm (относительная частота). Во второй половине десятилетия по газете «Известия» ipm слов с корнем «-*сталин*-» достигает 5680,64, что выше самого частотного существительного по новому частотному словарю русской лексики (сущ. «год», ipm 3727,5), ipm слов с корнем «-*ленин*-» в тот же период и по той же газете составляет 4178,02, что также выше ipm самого частотного существительного.

Самыми частотными словоформами оказываются имена собственные «Сталин» (227, «Известия», вторая половина) и «Ленин» (162, «Известия», вторая половина) в именительном падеже.

Одним из наиболее частотных производных слов оказывается прилагательное «сталинский» (104 употребления, «Известия», вторая половина).

Таблица 3. Фрагмент списка словоформ с корнем «-*сталин*-» с указанием абсолютной частоты

Известия, с 1930 по 1934 гг.		**Известия, с 1935 по 1939 гг.**	
Сталин	30	Сталин	227
Сталину	5	сталинский	104
сталинский	2	Ленина-Сталина	44
Сталино	2	Сталину	29
		Сталиным	19
		Сталино	5
		ленинско-сталинской	4
		ленинско-сталинского	4
		Маркса-Энгельса-Ленина-Сталина	3
		ленинско-сталинскую	3
Правда, с 1930 по 1934 гг.		**Правда, с 1935 по 1939 гг.**	
Сталин	38	Сталин	93
сталинский	11	сталинский	38
Сталину	7	Ленина-Сталина	24
сталинградского	4	Сталину	14
сталинградский	2	Сталиным	10
Сталинград	2	Сталино	5
Ленина-Сталина	2	ленинско-сталинской	4
Сталиным	1	Маркса-Энгельса-Ленина-Сталина	2
Сталино	1	ленинско-сталинская	2
Сталинградстроя	1	Сталинуголь	1
сталинградскому	1	Сталинград	1
Сталинграда	1	Сталинабад	1

Произвольно взяв одну из грамматических форм этого прилагательного, а именно, именительный падеж множественного числа, мы использовали функцию NGram Viewer «подстановочный знак», построив графики частотного

поведения десяти самых частотных биграмм с этой словоформой и существительными в правой позиции (рис.1–3).

Таблица 4. Фрагмент списка словоформ с корнем «ленин» с указанием абсолютной частоты

Известия, с 1930 по 1934		**Известия, с 1935 по 1939**	
Ленин	36	Ленин	162
ленинский	28	Ленина - Сталина	44
ленинградский	27	ленинградский	31
Ленинград	23	ленинский	27
ленинизм	3	Ленинград	25
марксистско-ленинской	2	ленинец	7
ленинградец	2	большевика-ленинца	7
марксистко-ленинскую	1	ленинско-сталинской	4
ленинградцы	1	ленинско-сталинского	4
		Маркса-Энгельса-Ленина-Сталина	3
		ленинско-сталинскую	3
		Ленина-	3
		большевик-ленинец	3
		по-ленински	2
		Маркса-Энгельса-Ленина	2
		ленинско-сталинская	2
		ленинградец	2
		чекист-ленинец	1
Правда, с 1930 по 1934		**Правда, с 1935 по 1939**	
Ленин	43	Ленин	62
ленинский	35	Ленина-Сталина	24
Ленинград	31	ленинский	16
ленинградский	29	Ленинград	10
ленинизм	7	ленинградский	6
марксизма-ленинизма	2	ленинско-сталинской	4
Ленинске	2	Маркса-Энгельса-Ленина-Сталина	2
Ленина-Сталина	2	ленинско-сталинская	2
марксизм-ленинизм	1	ленинец	2
ленинского	1	марксистско-ленинской	1
ленинским	1	ленинско-сталинского	1
антиленинских	1	ленинско-сталинский	1
		ленинско-кузнецких	1
		ленинская	1

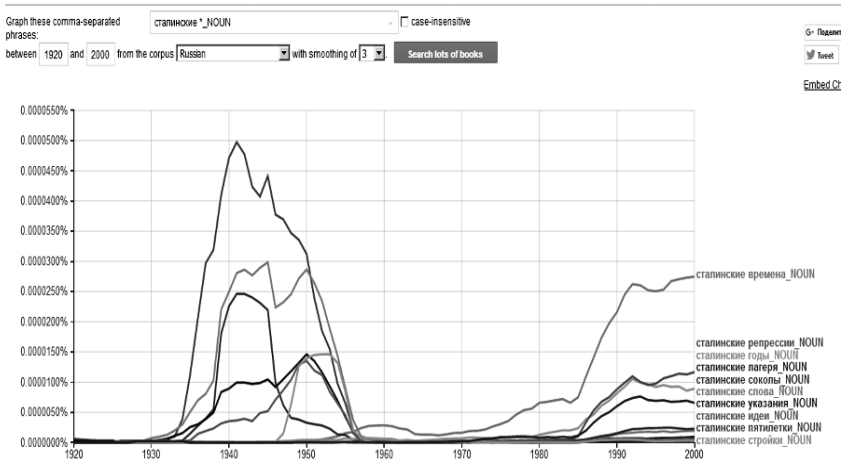


Рис. 1. Частотное поведение биграмм со словоформой «сталинские» в левой позиции с 1920 по 2000 гг.

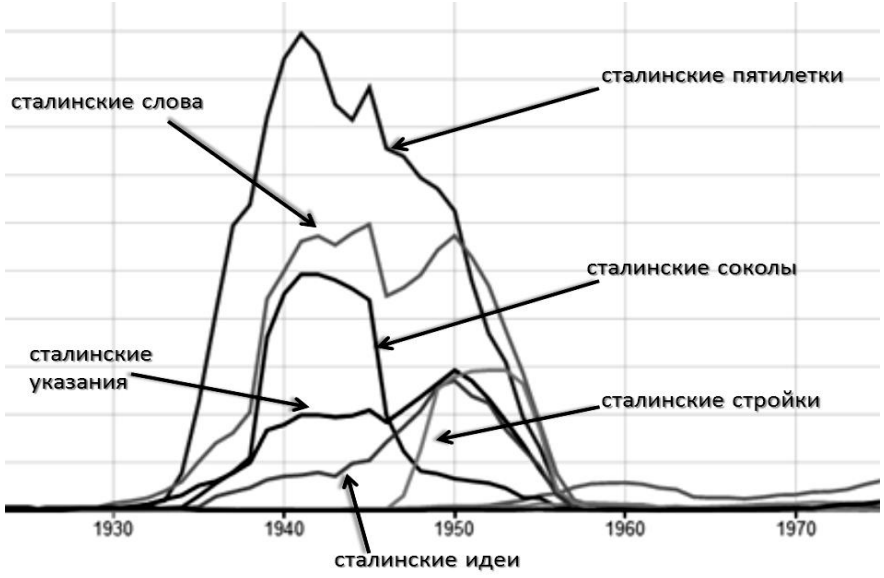


Рис. 2. Частотное поведение биграмм со словоформой «сталинские» в левой позиции с 1930 по 1975 гг.

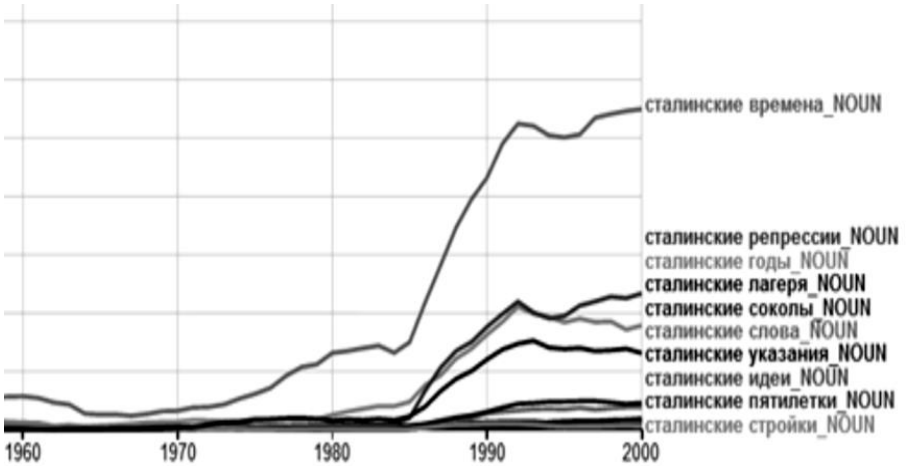


Рис.3. Частотное поведение биграмм со словоформой «сталинские» в левой позиции с 1960 по 2000 гг.

На рис. 1 видно вполне закономерное и не раз описанное нами явление: высокий частоты употребления лексики характерной для периода сталинского правления с 1930-х по 1950-е гг., резкое снижение её в середине 1950-х, и подъём в середине 1980-х гг.

В данном случае мы имеем дело с выраженным изменением коннотации от предельно позитивной до предельно негативной.

На рис. 2 видно, что с начала 1930-х до середины 1950-х гг. отмечается высокий уровень встречаемости словосочетаний, содержащих прилагательное «сталинские», которые являются пропагандистскими клише. Прилагательное «сталинские» имело в контексте этого временного периода выражено позитивную коннотацию.

Начиная с середины 1980-х гг., снова наблюдается рост сочетаний с этим прилагательным, как это видно на рис. 3, но более высокой становится частотность словосочетаний с негативной коннотацией.

На рис. 1 также видно, что частотность негативных словосочетаний в 1990-е годы значительно ниже, чем позитивных в 1930–1940-е. На основании этого, пожалуй, можно сделать предположение, что пропаганда сталинского режима велась интенсивней, чем его критика. При этом словосочетания с негативными коннотациями сами по себе носят скорее нейтральный характер («сталинские времена», «сталинские годы»). Такое экспрессивное выражение, как «сталинские преступления», даже не вошло в набор из десяти наиболее частотных биграмм и по частоте встречаемости значительно уступает словосочетанию «сталинские пятилетки» (рис.4).

Интерес, на наш взгляд, представляет и то обстоятельство, что коннотации варьируются не только в ходе исторического процесса, но в синхронических слоях лексики.

В текущий исторический период мы можем встретить текст как с негативным, так и с позитивным отношением к этому историческому лицу. Например: «*За XXI – бурный XXII съезд, когда Хрущева занесло, и он, оторвавшись от текста доклада, рассказывал о таких преступлениях Сталина, вскрывал такие нарывы и обнажал такие язвы, что все даже не появилось в печати – побоялись*». Михаил Козаков. Актерская книга (1978-1995).

«*Ельцин, разгромивший гигантскую державу из десяти часовых поясов, развеявший на сквозняках истории тысячелетние труды предков, — Ивана Калиты, Алексея Михайловича, Петра Первого, Екатерины Великой, Муравьева-Амурского, Семёнова-Тяньшаньского, Сталина, Фрунзе, космического Гагарина и ядерного Курчатова, — такого метафизического предательства не знали народы и царства*». Александр Проханов. Антисоветский Путин (2003) // «Завтра», 2003.08.13

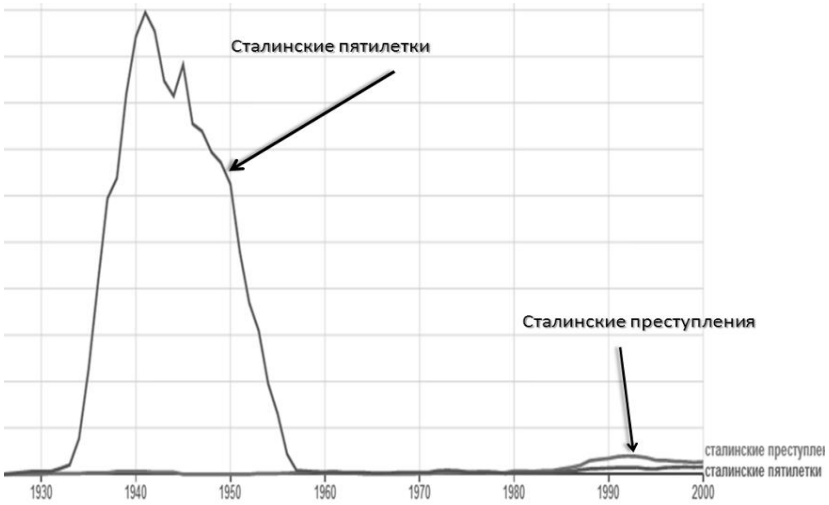


Рис. 4. Сопоставление частотного поведения кривых словосочетаний «сталинские пятилетки» и «сталинские преступления»

3. Слово «товарищ» — маркер коннотации

На рис. 5 видно, что слово «товарищ» в данном случае является маркером позитивной коннотации. Это же следует из приведённых ниже примеров, взятых из базы данных Google books:

Пример 1. «*В годы пятилеток, по указанию товарища Сталина, на территории, находящейся между Волгой и Уралом, развернулись большие геолого-разведочные работы, в результате которых открыто «Второе Баку»*. Огонек, 1947 г.

Пример 2. *«По указанию Сталина среди членов семьи Аллилуевых были произведены аресты».* «Воля: журнал узников тоталитарных систем». 1995.

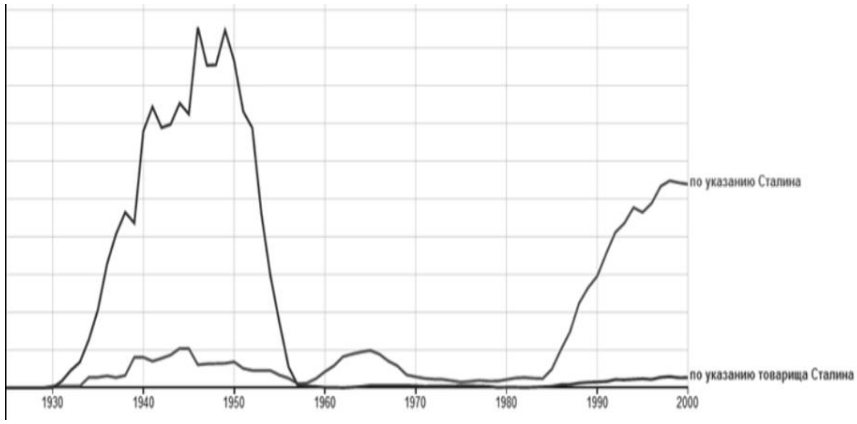


Рис. 5. Частотное поведение словосочетаний «по указанию Сталина» и по «указанию товарища Сталина»

4. Аплодисменты: грамматическая форма — контекст — отражение истории

Пример графика со словом «аплодисменты» приводился в публикации [9], Здесь мы рассмотрим некоторые новые аспекты этого необычного с точки зрения частотного поведения слова. Обратим внимание, что при задании условий построения графика система различала строчную и заглавную буквы (Это условие может задать пользователь).

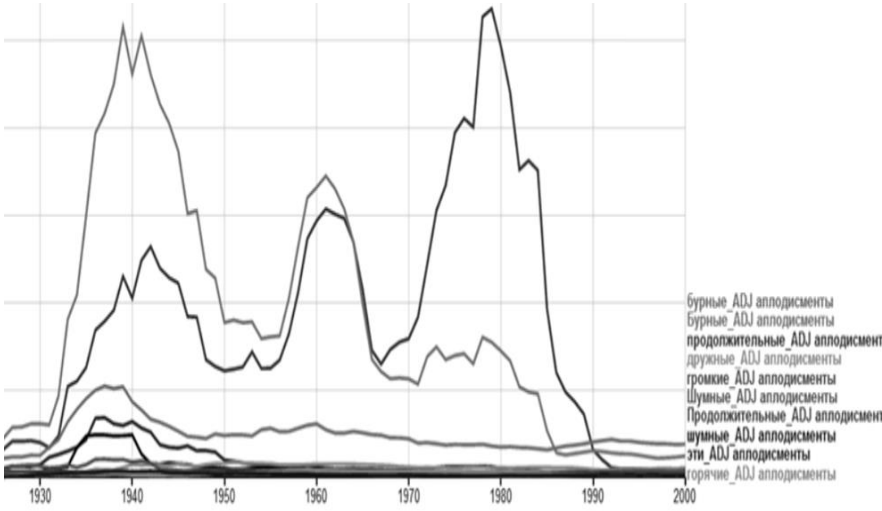
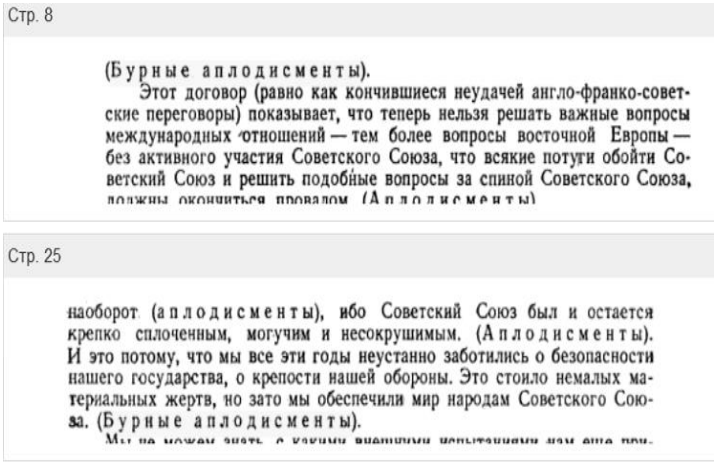


Рис.6. Частотное поведение биграмм со словоформой «аплодисменты» в правой позиции

Пример 3. Исторический журнал, Том 9,Выпуск 2 изд-во «Правда» 1939 г.



Стр. 8

(Бурные аплодисменты).

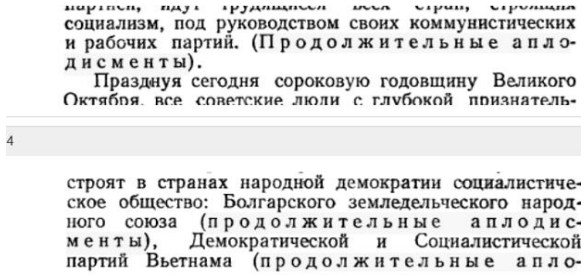
Этот договор (равно как кончившиеся неудачей англо-франко-советские переговоры) показывает, что теперь нельзя решать важные вопросы международных отношений — тем более вопросы восточной Европы — без активного участия Советского Союза, что всякие потуги обойти Советский Союз и решить подобные вопросы за спиной Советского Союза, должны окончиться провалом. (Аплодисменты).

Стр. 25

наоборот (аплодисменты), ибо Советский Союз был и остается крепко сплоченным, могучим и несокрушимым. (Аплодисменты). И это потому, что мы все эти годы неустанно заботились о безопасности нашего государства, о крепости нашей обороны. Это стоило немалых материальных жертв, но зато мы обеспечили мир народам Советского Союза. (Бурные аплодисменты).

Мы не можем знать, с какими внешними испытаниями нам еще при-

Пример 4. Н.С, Хрущев «40 лет Великой Октябрьской социалистической революции июль, 1957»



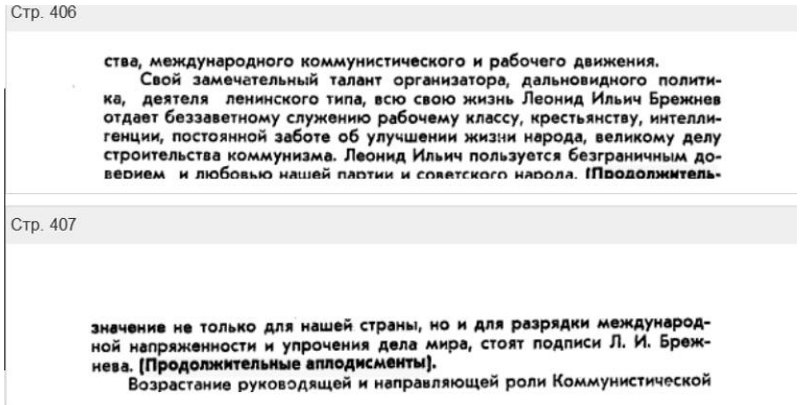
партией, идут трудящиеся всех стран, строящие социализм, под руководством своих коммунистических и рабочих партий. (Продолжительные аплодисменты).

Празднуя сегодня сороковую годовщину Великого Октября, все советские люди с глубокой признатель-

4

строят в странах народной демократии социалистическое общество: Болгарского земледельческого народного союза (продолжительные аплодисменты), Демократической и Социалистической партий Вьетнама (продолжительные апло-

Пример 5. Коммунист вооруженных сил 1977



Стр. 406

ства, международного коммунистического и рабочего движения.

Свой замечательный талант организатора, дальновидного политика, деятеля ленинского типа, всю свою жизнь Леонид Ильич Брежнев отдает беззаветному служению рабочему классу, крестьянству, интеллигенции, постоянной заботе об улучшении жизни народа, великому делу строительства коммунизма. Леонид Ильич пользуется безграничным доверием и любовью нашей партии и советского народа. (Продолжитель-

Стр. 407

значение не только для нашей страны, но и для разрядки международной напряженности и упрочения дела мира, стоят подписи Л. И. Брежнева. (Продолжительные аплодисменты).

Возрастание руководящей и направляющей роли Коммунистической

Наиболее частотными биграммами являются биграммы «Бурные аплодисменты» и «Продолжительные аплодисменты». Как видно из примеров 3–5, в приведённых текстах эти биграммы присутствуют в виде ремарок. Этим объясняется преобладание написания первого слова с заглавной буквы. Анализ результатов поиска показывает, что эти два словосочетания присутствуют в основном в публикциях стенограмм выступлений политических деятелей.

График на рис. 6 отличает странная симметричнсть конфигурации кривых, каждая из которых имеет три «пика». В 1939–1941 гг. наблюдается первый «пик с двумя «вершинами», В период с начала 1930-х гг. до середины 1950-х гг. преобладает словосочетание «Бурные аплодисменты». С середины 1950-х до середины 1960-х обе кривые почти совпадают, а середины 1960-х до начала 1990-х преобладает словосочетание «Продолжительные аплодисменты» с пиком кривой в 1980 году. По-видимому, клише «Бурные аплодисменты» в качестве ремарки должно было обозначать энтузиазм и эмоциональный подъем аудитории. Можно предположить, что данные графики корреспондируют со стилем работы трёх руководителей СССР: И.В. Сталина, Н.С. Хрущева и Л.И. Брежнева. Зто предположения подкрепляется частотным поведением словосочетания «переходящие в овацию» (рис.7).

Рис. 7. Частотное поведение триграммы «переходящие в овацию»

Пример 6. Стенографический отчет XVI съезда ВКП (б), 1930:

«*Председательствующий . Товарищи, т. Владимирский и т. Каганович отказались от заключительного слова. (Аплодисменты.) Заключительное слово имеет т. Сталин. (Продолжительные аплодисменты, переходящие в овацию)*».

Отметим ещё одну особенность этих графиков. На рис. 8 отчетливо видно, что частотность биграммы «продолжительные аплодисменты» в именительном падеже намного превосходит частотность биграмм в других падежах. Чуть выше частотность биграммы творительного падежа, что отражает частое

наличие в текстах-репортажах выражений типа «продолжительными аплодисментами встретили…»

Заключение

Мы попытались показать динамику изменений коннотаций слов и словосочетаний, выбрав три примера из имеющегося у нас обширного материала. Проявления и механизмы этих изменений различны. Во всех трех случаях изменения вызваны изменением общественно-политической ситуации, изменением отношения к определенному историческому лицу и эпохе.

В первом случае это выражается в появлении и/или росте частотности словосочетаний с явным или имплицированным негативным значением.

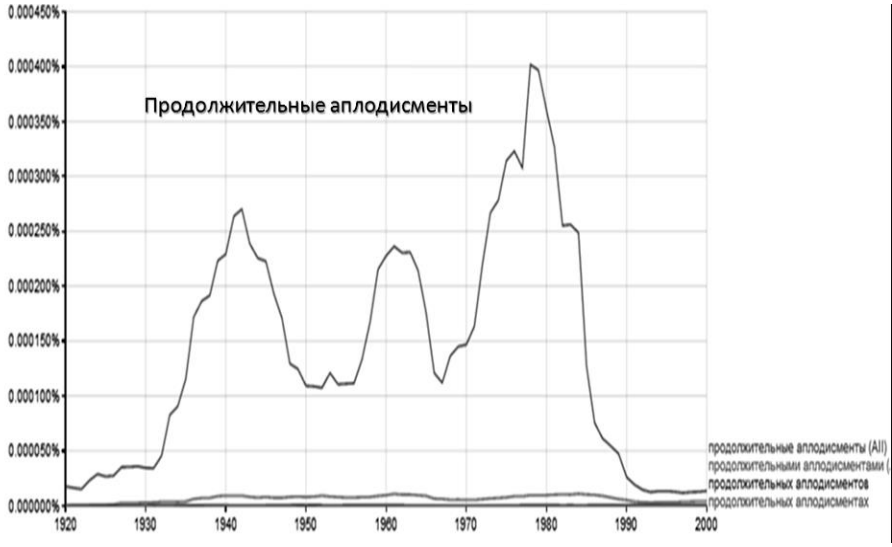


Рис.8. Сопоставление частотного поведения различных грамматических форм биграммы «продолжительные аплодисменты»

Во втором случае коннотацию определило маркирующее слово. Впрочем, иногда выражение «товарищ Сталин» может иметь ироническую окраску.

Наконец, третий случай показывает, как в ремарке к стенограмме может присутствовать семантический нюанс, который отражает некоторые изменения в протоколе политических собраний, типичных для советского времени. Заметим, что с начала 1990-х гг. частотность всех биграмм со словом аплодисменты резко падает.

Таким образом, диахронические исследования на базе корпусов позволяют выявлять тонкие семантические изменения в языке.

Литература

[1] Michel J-B. et al. Quantitative Analysis of Culture Using Millions of Digitized Books science. Science, 2011, 331, 176. DOI: 1126/Science.1199644. http://www.sciencemag.org/content/331/6014/176.full.html.

[2] Захаров В.П., Масевич А.Ц. Диахронические исследования на основе корпуса русских текстов Google Books Ngram Viewer // Структурная и прикладная лингвистика. Вып. 10. СПб., 2014. С. 303-327.

[3] Масевич А.Ц., Захаров В.П. Опыт корпусно-ориентированного историко-культурного исследования исторической и политической лексики // Библиосфера. 2016. № 2. С. 47-55.

[4] Масевич А. Ц. Захаров В. П. Методы корпусной лингвистики в исторических и культурологических исследованиях // Компьютерная лингвистика и вычислительные онтологии: сб. научн. статей. Труды XIX Международной объединенной научной конференции «Интернет и современное общество» (IMS-2016), Санкт-Петербург, 22 – 24 июня 2016 г. СПб: Университет ИТМО, 2016. С. 24-43.

[5] Масевич А.Ц., Захаров В.П. Методы корпусной лингвистики в исторических исследованиях. https://www.researchgate.net/publication/309386435_Metody_korpusnoj_lingvistiki_v_istoriceskih_issledovaniah_Corpus_Linguistics_Methods_in_Historical_Studies.

[6] Масевич А.Ц., Захаров В.П. Лингвистическая картина российской истории. https://www.researchgate.net/publication/309386514_Lingvisticeskaa_kartina_rossijskoj_istorii_Linguistic_portrait_of_20th_century_Russian_history_a_corpora-based_study.

[7] Kilgarriff A., Rychly P., Smrz P., Tugwell D. The Sketch Engine. In Proc. EURALEX 2004, Lorient, France. P. 105–116.

[8] Вахтин Н. Б. Социолингвистика и социология языка : учеб. пособие / Н. Б. Вахтин, Е. В. Головко. СПб. : Изд. Центр "Гуманитар. Акад." : Европ. ун-т в СПб., 2004. 335 с.

[9] Захаров В. П. Сочетаемость через призму корпусов // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 27-30 мая 2015 г.). Вып. 14 (21). М., Изд-во РГГУ, 2015. Т. 1. С. 667-682.

Semantic transformation of political vocabulary

A.Ts.Masevich [1], V.P. Zakharov[2]

[1] St.-Petersburg State Institute of Culture

[2] St.-Petersburg State University

The paper presents results of diachronic research of the frequency behavior of three items of Russian political vocabulary. As tools, we used the Google Books Ngram Viewer to create a set of charts of the frequency behavior of three selected items in the period from 1920 to 2000. This instrument creates charts of changing frequency of lexical units — words and collocation — within given period. In addition we used the corpus management system Sketch Engine to create corpora of texts from newspapers “Pravda” and “Izvestiya” covered period from 1930 to 1939. Various statistical characteristics were obtained from these data. We found that within a given chronological period connotations of the items have changed. We revealed certain correlations with the historical process.

Keywords: corpora of texts, diachronic studies, frequency of lexical units, models of words behavior, semantic field.