

Исследование в НКРЯ

ЧАСТЬ 1

Я провела исследования в Национальном корпусе русского языка с целью выяснить, насколько менялась с течением времени частота употреблений слов, которые с точки зрения современного языка принято считать синонимами – «НЫНЕ» и «СЕЙЧАС». Однако важно заметить, что стилистическая окраска данных слов не совпадает ввиду различий в истории их употребления. Именно по этой причине данное исследование весьма и весьма любопытно с лингвистической точки зрения.

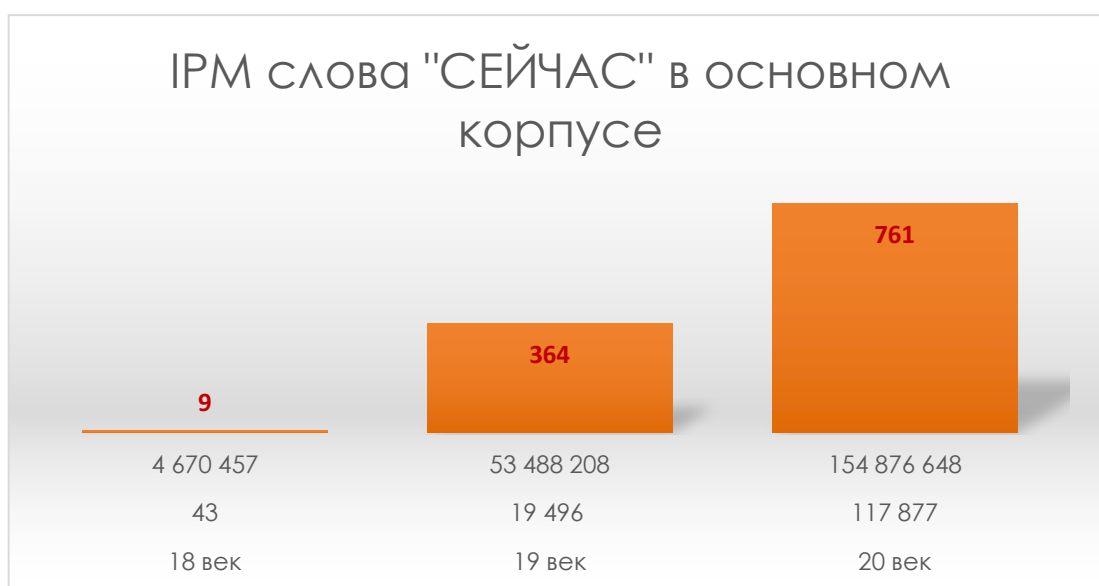
Проведено оно было в основном и поэтическом корпусах, важно было сравнить и сопоставить частоты употреблений обеих лексических единиц в XVIII, XIX и XX веках.

Сравнение я проводила, исходя из IPM (количество вхождений, деленное на общее количество слов и умноженное на миллион) каждого слова на трех заданных временных промежутках.

Внизу приведены диаграммы, иллюстрирующие полученные мной результаты исследования в основном корпусе.

(1ая строчка внизу под столбиками - общее количество слов
2ая строчка внизу под столбиками – количество вхождений)

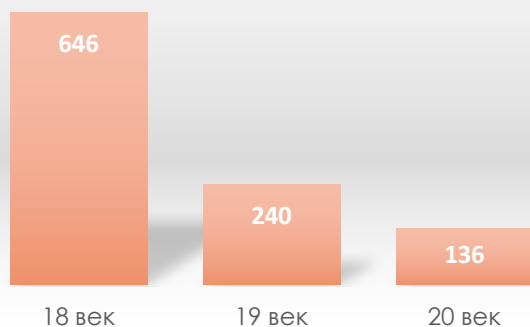
*Для удобства IPM округлен до десятых



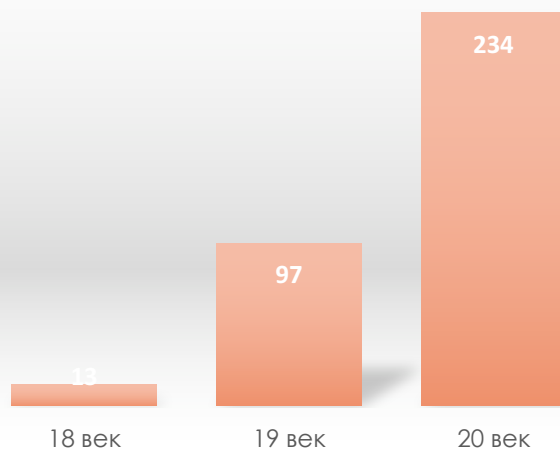
На представленных диаграммах отчетливо прослеживается разница в статистике употребления слов «НЫНЕ» и «СЕЙЧАС» на одних и тех же временных промежутках. Интересно, что в то время как частота употребления первого слова с годами становится всё меньше и меньше, частота употребления более привычного нам второго слова идет по полностью противоположному пути. Таким образом, вследствие потери ей своей актуальности, справедливо будет отнести лексическую единицу «НЫНЕ» к так называемым «устаревшим словам».

Ниже представлены диаграммы, иллюстрирующие статистику для поэтического корпуса:

IPM слова "НЫНЕ" в поэтическом корпусе



IPM слова "СЕЙЧАС" в поэтическом корпусе



Важно отметить, что несмотря на то, что тенденции увеличения частоты употребления слова «СЕЙЧАС» и уменьшения частоты употребления слова «НЫНЕ» совпадают в основном и поэтическом корпусах, изначальное количество данных в самих корпусах различно:

	A	B	C	D
1	СЛОВО "НЫНЕ"	КОЛИЧЕСТВО ВХОЖДЕНИЙ	ВСЕГО СЛОВ	IPM
2	18 век	2 950	4 670 457	632
3	19 век	5 975	53 488 208	112
4	20 век	10 346	154 876 648	67
5				
6				
7				
8	СЛОВО "СЕЙЧАС"	КОЛИЧЕСТВО ВХОЖДЕНИЙ	ВСЕГО СЛОВ	IPM
9	18 век	43	4 670 457	9
10	19 век	19 496	53 488 208	364
11	20 век	117 877	154 876 648	761
12				
13				
14				
15	СЛОВО "НЫНЕ"	КОЛИЧЕСТВО ВХОЖДЕНИЙ	ВСЕГО СЛОВ	IPM
16	18 век	596	921 424	646
17	19 век	858	3 572 009	240
18	20 век	889	6 501 047	136
19				
20				
21				
22	СЛОВО "СЕЙЧАС"	КОЛИЧЕСТВО ВХОЖДЕНИЙ	ВСЕГО СЛОВ	IPM
23	18 век	12	921 424	13
24	19 век	347	3 572 009	97
25	20 век	1 526	6 501 047	234

ОСНОВНОЙ КОРПУС

ПОЭТИЧЕСКИЙ КОРПУС

Количество вхождений также различно – куда большим оно является в основном корпусе. Вероятно, здесь дело прежде всего в том, что оба данных слова с лингвистической точки зрения скорее свойственны нехудожественным текстам.

Так или иначе, вне зависимости от количества исходных данных в корпусах, в отношении двух представленных слов прослеживается определенная тенденция.

ЧАСТЬ 2

Далее мне необходимо было узнать, возможны ли были во времена правления Екатерины Второй следующие диалоги:

— *Отдай!*

— *Пиши пропало. Что пряжка, третьего дня эта поганка у меня с груди звезду Александра Невского уперла! Любит, сволочь, блестящее.*

— *Ох, ну просто немыслимо! Нет, как можно быть таким лопухом?!*

— *Я гнева вашего никак не растолкую. Он в доме здесь живет, великая напасть!*

Годы правления Екатерины Второй: 1762-1796. Таким образом, зная годы правления императрицы, мы можем по наиболее характерным словам и выражениям проверить данные диалоги на типичность для тех времен. Рассмотрим первый диалог:

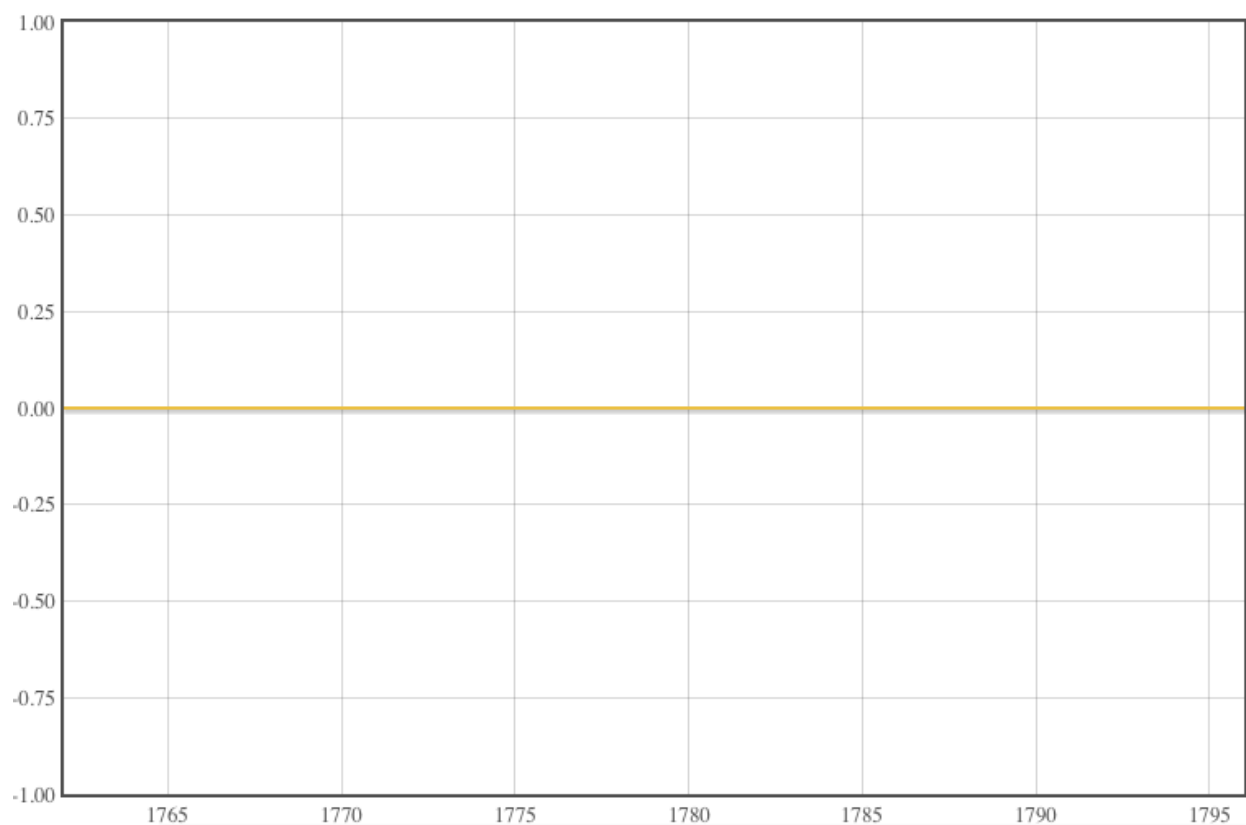
— *Отдай!*

— *Пиши пропало. Что пряжка, третьего дня эта поганка у меня с груди звезду Александра Невского уперла! Любит, сволочь, блестящее.*

Возьмем выражение «ПИШИ ПРОПАЛО» и скопируем в поиск в основном корпусе. В графе «распределение по годам» зададим построение графика на период от 1762 до 1796 года. Получаем следующий результат:

Годы от 1762 до 1796 со сглаживанием 3

построить



Таким образом, нетипичность первого диалога для годов правления Екатерины Второй очевидна.

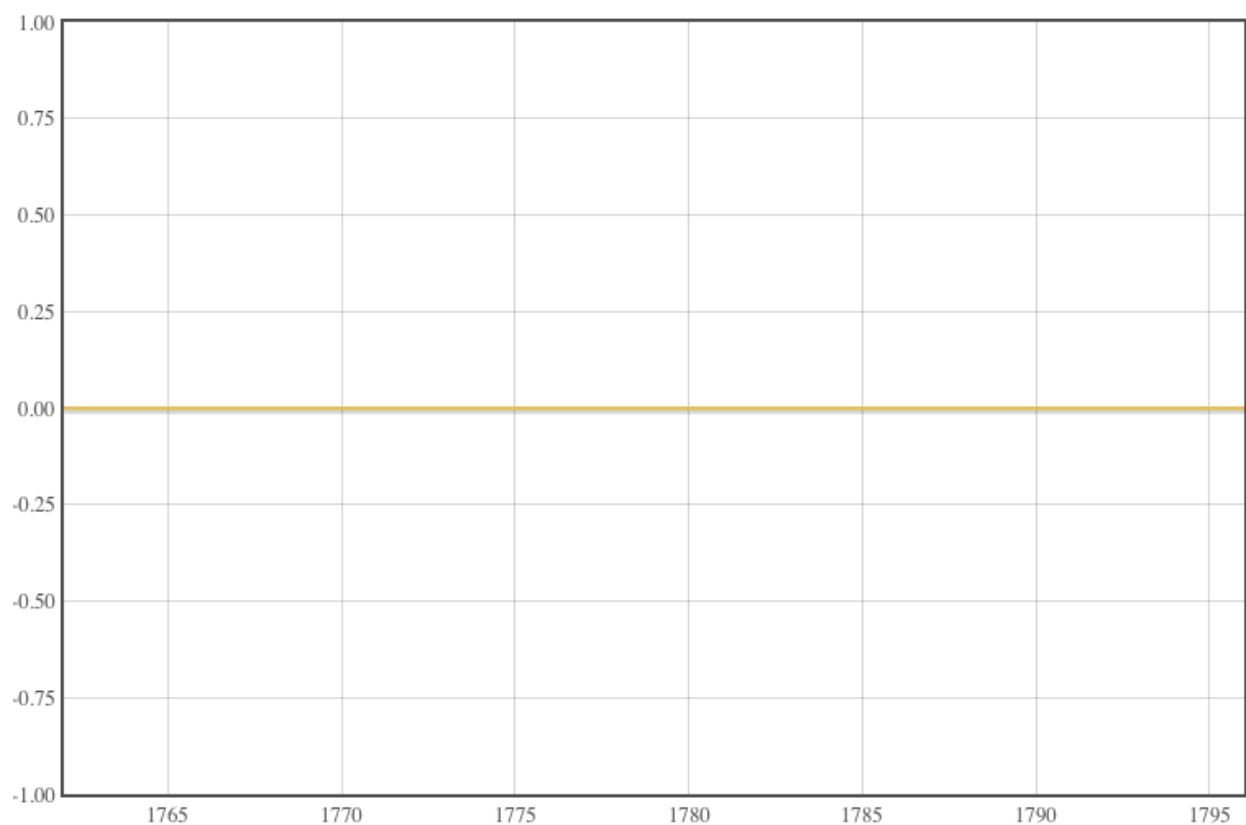
Аналогично, используя слово «НЕМЫСЛИМО», установим, возможен ли был второй диалог на том же временном интервале.

— Ох, ну просто невозможно! Нет, как можно быть таким лопухом?!

Получаем следующий график:

Годы от 1762 до 1796 со сглаживанием 3

построить



Из чего делаем вывод о невозможности подобного диалога в указанный период.

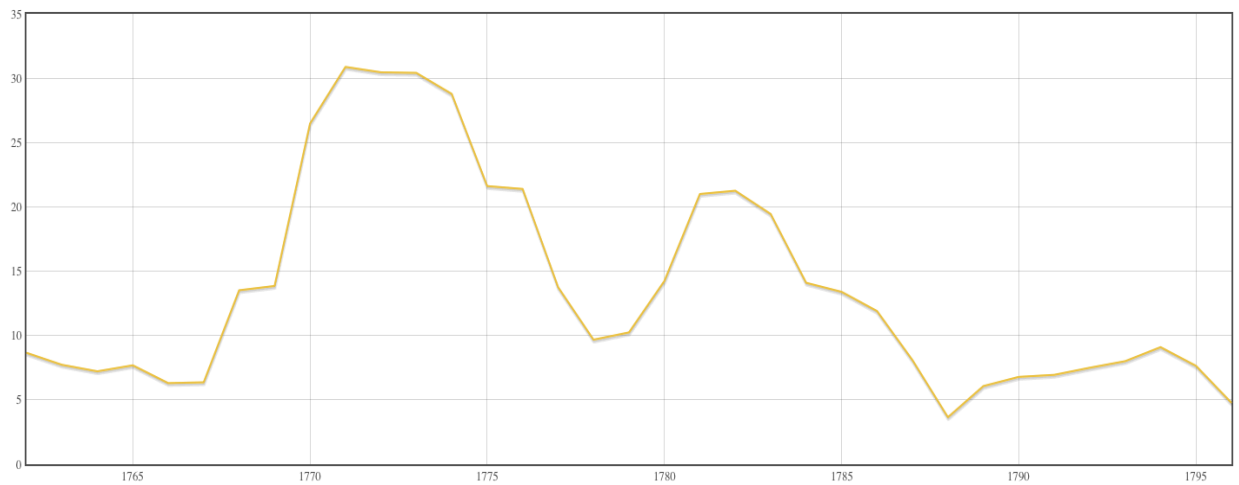
И, наконец, третий диалог:

— Я гнева вашего никак не растолкую. Он в доме здесь живет, великая напасть!

Для слов «РАСТОЛКОВАТЬ» и «НАПАСТЬ» на том же временном промежутке получаем следующие графики:

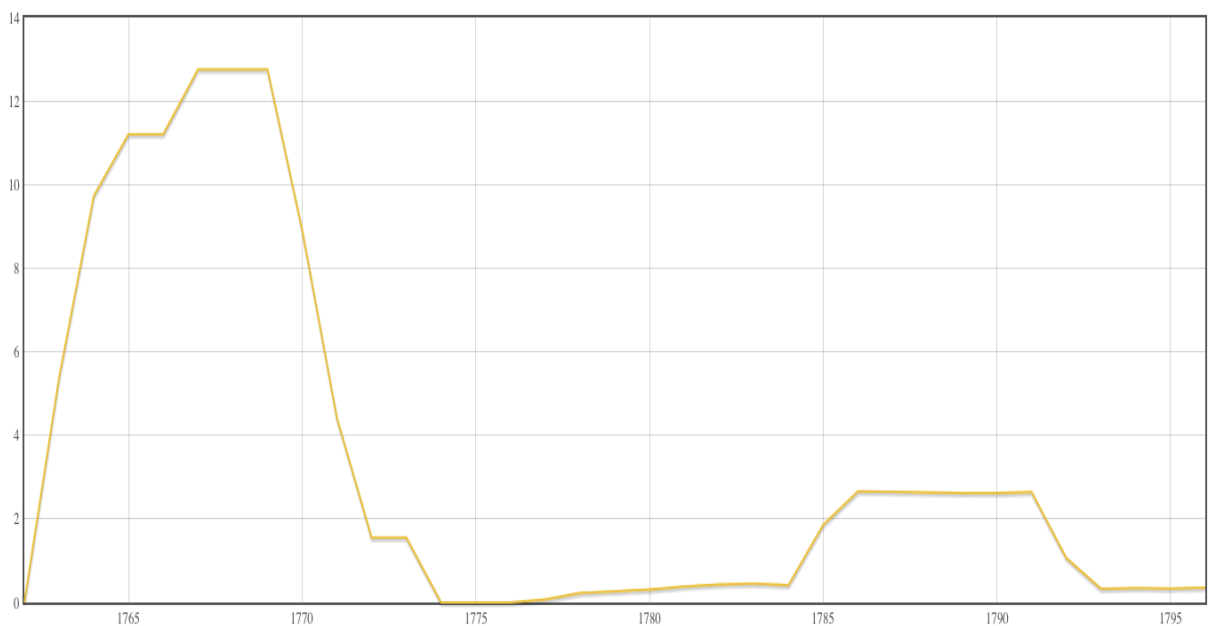
напасть

Годы от 1762 до 1796 со сглаживанием 3 построить



растолковать

Годы от 1762 до 1796 со сглаживанием 3 построить



Получить файл

Можем сделать вывод о явной вероятности данного диалога для екатерининских времен.

ЧАСТЬ 3

Следующим пунктом рассмотрим динамику употребления слова «щас» на период с 1950 по 2018 год.

На представленном ниже графике видно, что пик частоты употребления данного слова приходится на 2008 год. Далее виден существенный спад, объясняемый потерей данной жаргонной словоформы своей актуальности.

Слова или сочетания слов, через запятую: щас

Годы с 1950 по 2018 со сглаживанием 3

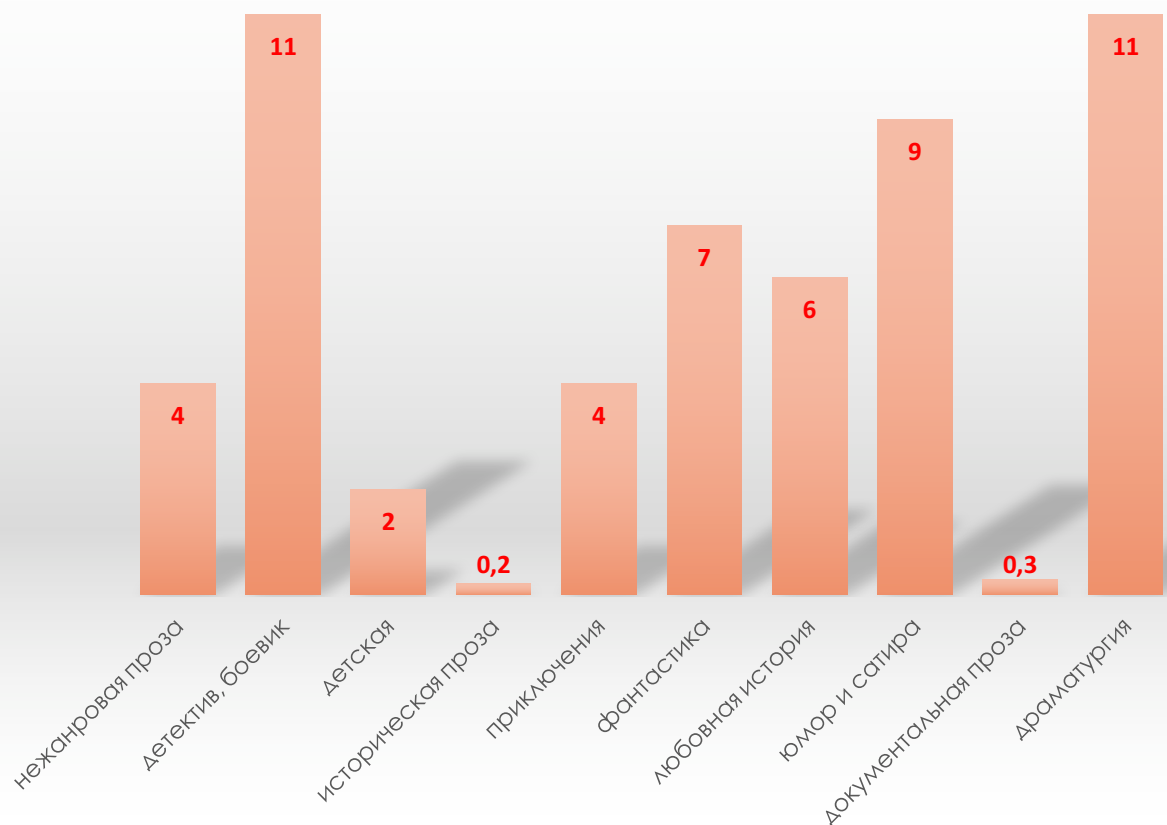
построить



Теперь рассмотрим IPM слова «ЩАС» с точки зрения различных литературных жанров. Весьма предсказуемо, что для таких жанров, как, например, историческая или документальная проза слово совершенно не характерно. Дело в первую очередь в его нелитературности, «низком», бытовом стиле. Подобные слова хоть и встречаются в литературе разных жанров, но скорее являют собой средство придания реалистичности описанному.

ЖАНР	КОЛИЧЕСТВО ВХОЖДЕНИЙ	ВСЕГО СЛОВ	IPM
нежанровая проза	283	69 230 465	4
детектив, боевик	85	7 587 093	11
детская	10	5 868 568	2
историческая проза	3	10 367 857	0,2
приключения	15	3 610 684	4
фантастика	55	8 199 359	7
любовная история	6	973 982	6
юмор и сатира	43	4 595 212	9
документальная проза	3	10 808 635	0,3
драматургия	36	3 378 029	11

IPM слова "ЩАС" для различных жанров



ЧАСТЬ 4

Я проанализировала выдачу по запросу «язык+ существительное в родительном падеже» и выяснила, что наиболее часто встречающееся словосочетание такой грамматической конструкции – «язык программирования». Не стану скрывать, результат меня удивил, так как я привыкла смотреть на язык с точки зрения его национальной принадлежности и культурных особенностей. Однако данная статистика вполне объяснима. В век технического прогресса язык (в привычном нам понимании этого слова) теряет свою значимость, уступая место более прогрессивному и многоаспектному компьютерному языку. Сложно оценить данное явление с точки зрения его положительности или отрицательности, однако реальность такова, что язык программирования действительно становится все более важным как для общения, так и для более сложных по своей сути процессов.