

**А.В. Кашкина**

*Воронежский государственный университет, г. Воронеж*

**НАИБОЛЕЕ УПОТРЕБИТЕЛЬНЫЕ МАРКЕМЫ И ИХ  
ЗАКОНОМЕРНОСТИ  
THE MOST FREQUENT MARKEMES AND THEIR REGULARITIES**

*Ключевые слова: маркама, маркемный анализ, индекс тематической маркированности, частота, русская поэзия*

*Keywords: markeme, markemic analysis, Index of Thematic Markedness, frequency, Russian poetry*

Задачи обработки и анализа крупных текстовых массивов с целью выявления общих закономерностей привели к разработке в лингвистике различных квантитативных методик, основанных на применении компьютерных программ. Один из таких методов – маркемология, или маркемный анализ, разработанный проф. А.А. Кретовым.

Маркемный анализ базируется на понятиях индекса тематической маркированности и маркамы. Индекс тематической маркированности (ИнТеМ) – числовой параметр, отражающий роль той или иной словоформы в тексте, вычисляемый по формуле

$$\text{ИнТеМ} = \text{Ч-вес} - \text{Ф-вес},$$

где Ч-вес – относительная частота словоформы в данном тексте, Ф-вес – функциональный вес словоформы, зависящий от ее длины в звуках (Кретов 2008).

Маркемы – словоформы текста, обладающие положительным ИнТеМ и удовлетворяющие ряду критериев:

1) грамматические: имя существительное, в форме именительного падежа, единственного числа (кроме *pluralia tantum* и слов типа *глаза*), неодушевленное (кроме слова *человек*).

2) лексико-семантические: имя нарицательное, не артефакт (кроме артефактов с символическим значением), не классификатор (как *множество* в сочетаниях типа *множество людей, книга*), не название месяца/дня недели, не связано с каким-либо литературным жанром (*действие, явление* в пьесах) и не является его названием.

3) лексико-стилистические: не окрашено стилистически.

4) лексико-синтаксические: не является обращением, вводным словом (Кретов 2010, Кашкина 2012).

Метод маркемного анализа применялся нами для изучения языка русской поэзии XVIII-XX вв. на материале творчества 105 авторов, распределенных по 7 хронологическим срезам (XVIII в., начало, середина и конец XIX в., начало XX в., 20-50-е гг. и 60-80-е гг. XX в.). Во всех срезах, кроме начала XX в. (16 авторов) рассматриваются произведения 15 авторов; творчество А.Н. Плещеева рассматривается в 2 периодах (середина и конец XIX в.). Маркемный список для каждого из исследованных авторов включает

50 маркем, выделенных в соответствии с вышеизложенными критериями и обладающих наибольшим ИнТеМ (Кашкина 2013: 2-3). Для автоматического определения ИнТеМ маркем и ранжирования словоформ по убыванию этого параметра используется программа ТемАЛ – «Тематический анализатор лексики», созданная И.В. Поповой под руководством проф. А.А. Кретьова и И.Е. Ворониной (Кретьов, Воронина, Попова 2010).

Маркемы какого-либо текста, на основании их определения, имеют в данном тексте частоту, превышающую средние для данного языка показатели частотности словоформ аналогичной длины. Однако в маркемных списках исследованных нами русских поэтов присутствуют маркемы, частота которых значительно (в десятки раз) превышает частоту других маркем. Так, у Г.Р. Державина маркема *сердце* встречается 47 раз, тогда как большинство прочих элементов его маркемного списка – в среднем 10 раз. Любопытно, что набор таких маркем сходен у большинства рассмотренных нами поэтов. В настоящей статье мы рассмотрим наиболее употребительные маркемы и связанные с ними закономерности.

Встречаются маркемы, являющиеся высокоупотребительными лишь у небольшого количества авторов в одном хронологическом срезе. Так, для XVIII в. это *бытие, прах, добродетель*. Исключительная частота таких маркем обуславливается как спецификой того или иного временного промежутка (в меньшей степени), так и особенностями творчества отдельного автора.

Велико число маркем, встречающихся в качестве высокочастотных в двух последовательных срезах. Так, *премудрость, слава, блеск, страсть, страх, сладость*, обладают высокой частотой у ряда авторов XVIII и первой трети XIX в. Возможна и разрывная высокоупотребительность: *вечность* – в XVIII и начале XX в., *молодость* – в начале XIX и в 20-50-х гг. XX в. и т.п. Гораздо больший интерес с точки зрения выявления общих особенностей русского поэтического языка представляют маркемы, являющиеся высокоупотребительными у представителей всех или большинства хронологических срезов. В таблице 1 представлены такие маркемы, отмеченные в пяти из семи исследуемых срезов. Числа обозначают отношение количества авторов, у которых эта маркема исключительно частотна, к их общему количеству в срезе. Эти данные используются и для построения графиков, отражающих динамику частоты попадания маркем в наиболее употребительные (см. рис. 1-2).

Таблица 1. Маркемы, высокочастотные у представителей 5 хронологических срезов

Срез	<i>радость</i>	<i>воздух</i>	<i>свет</i>	<i>слово</i>	<i>взгляд</i>	<i>душа</i>	<i>ветер</i>	<i>земля</i>
XVIII	0,266	0,266	0,266	0	0	0	0	0
1-XIX	0,666	0	0	0,133	0,133	0,133	0	0
2-XIX	0,266	0,066	0,066	0,133	0,333	0,133	0,133	0,066
3-XIX	0	0	0,133	0,066	0,333	0,266	0,133	0,133
1-XX	0,4375	0,25	0	0	0,1875	0	0,25	0,1875

2-XX	0	0,066	0,333	0,133	0	0,066	0,333	0,333
3-XX	0,066	0,133	0,333	0,066	0,266	0,2	0,4	0,266

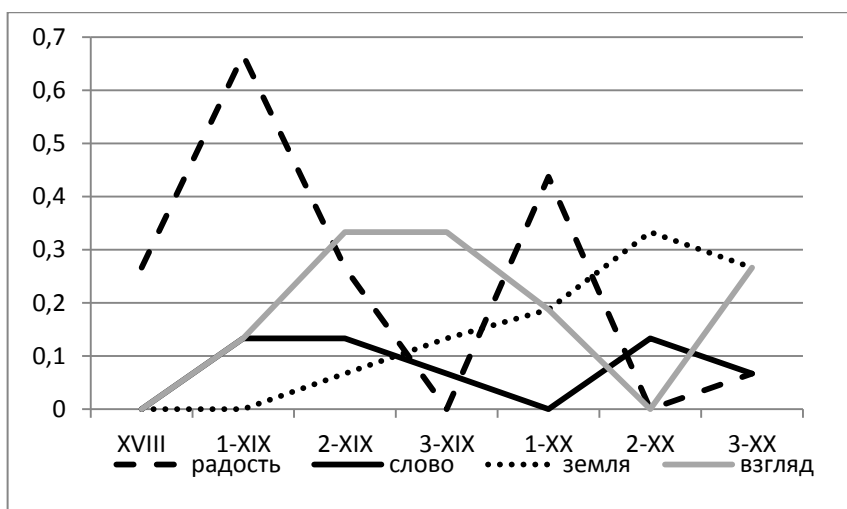


Рис. 1. Маркемы *радость*, *слово*, *земля*, *взгляд*: динамика

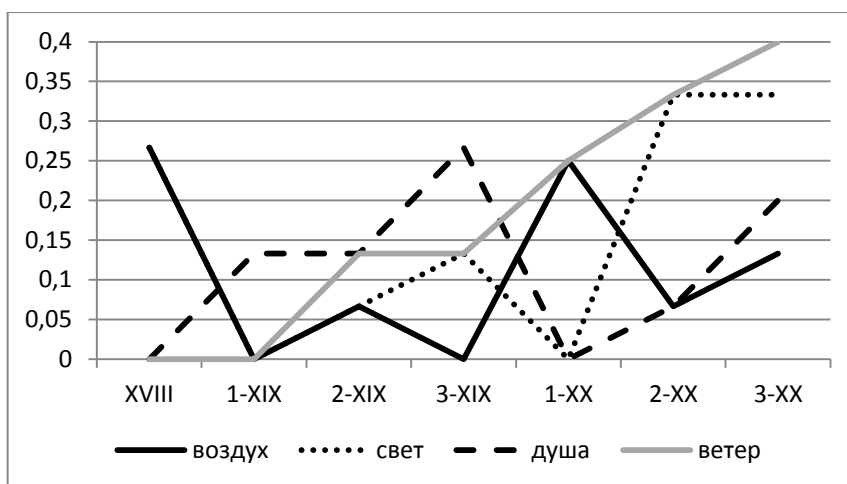


Рис. 2. Маркемы *воздух*, *свет*, *душа*, *ветер*: динамика

Наиболее типичными в роли высокоупотребительных для XVIII в. оказываются *радость*, *воздух*, *свет*. *Радость* достигает пика своей распространенности в первой трети XIX в. (является высокочастотной более чем у половины авторов), что может обуславливаться важностью маркемной лексики, описывающей эмоциональное состояние, для этого периода. В этом хронологическом срезе такие маркемы, как *душа*, *взгляд*, *слово*, оказываются высокоупотребительными лишь у 2 авторов. *Взгляд* обладает максимальной распространенностью (у 5 авторов) во 2-последней третях XIX в. В последней трети XIX в. соответствующего максимума достигает маркема *душа* (3 автора). *Радость*, напротив, в этом хронологическом срезе не является высокочастотной ни у одного из авторов, что может быть связано с общей пессимистичностью, мрачной окраской поэтических произведений этого времени. В начале XX в. пика своей распространенности достигает *воздух* (у 4 авторов), для маркемы *радость* это вторая точка экстремума (здесь она является высокочастотной почти у половины авторов). В 20-50-х

гг. XX в. наибольшее для них значение распространенности приобретают маркемы *свет, земля* (у 5 авторов). Они соответствуют двум преобладающим в этом периоде маркемным группам (Кашкина 2011: 98) – маркемам, обозначающим природные явления и объекты (так называемые «натурфакты») и маркемам, описывающим чувственно воспринимаемые характеристики окружающей действительности (визуальные, как *свет*, аудиальные: *звон, шелест* и прочие). Маркема *ветер* чаще всего выступает в роли высокоупотребительной в 60-80-х гг. XX в. Другие типичные в этом срезе высокочастотные маркемы – *свет, взгляд, земля*.

Число маркем, оказывающихся высокоупотребительными у представителей 6 из 7 проанализированных нами хронологических срезов, невелико – 3 маркемы (*человек, народ, жизнь*, см. табл. 2, рис. 3).

Таблица 2. Маркемы, высокоупотребительные у представителей 6 срезов

Срез	<i>человек</i>	<i>народ</i>	<i>жизнь</i>
XVIII	0,466	0,2	0,133
1-XIX	0,266	0,133	0,133
2-XIX	0,533	0,266	0,266
3-XIX	0	0,2	0,6
1-XX	0,4375	0	0
2-XX	0,333	0,133	0,466
3-XX	0,666	0,066	0,6

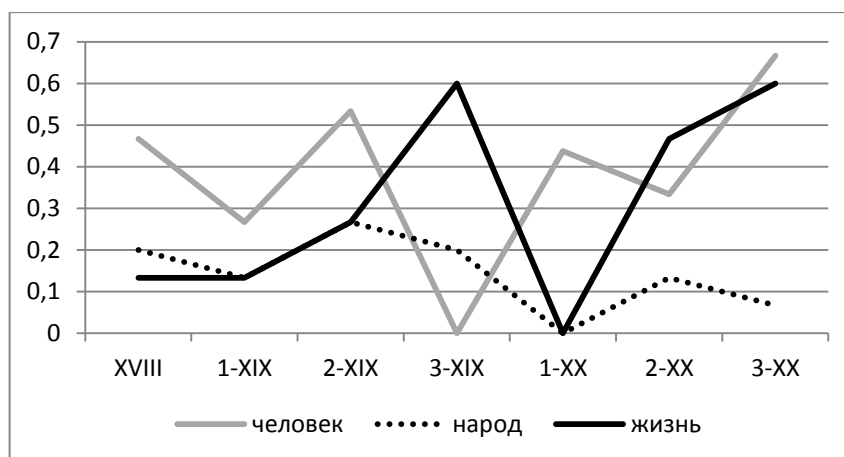


Рис. 3. Маркемы *человек, народ, жизнь*: динамика

Распространенность этих маркем (*человек, жизнь, народ*) значительно изменяется по срезам. Так, в XVIII в. отмечены все эти маркемы, но наиболее типичной из них является маркема *человек*, оказывающаяся высокоупотребительной почти у половины авторов. Затем, в первой трети XIX в. количество авторов, у которых маркемы *человек, народ* оказываются высокочастотными, снижается. Во второй трети XIX в. повышается количество авторов, употребляющих все три маркемы. Максимум своей распространенности в качестве высокочастотной (у 4 авторов) достигает маркема *народ*, как маркема, характерная для гражданской поэзии. В

последней трети XIX в. центральное место занимает *жизнь*, являющаяся высокоупотребительной у 9 из 15 авторов. Маркема *человек* в этом хронологическом срезе высокой частотностью ни у одного из поэтов не обладает. В начале XX в. картина меняется – маркема *человек* является одной из наиболее употребительных в произведениях 7 представителей этого хронологического среза, тогда как *жизнь* и *народ* в число высокочастотных не попадают. В 20-50-х гг. XX в. наиболее высокое положение среди представленных в данной выборке занимает маркема *жизнь*, часто употребляемая в произведениях почти половины авторов. Это также период второго пика распространенности маркемы *народ*, что связано с патриотической тематикой произведений. В 60-80-х гг. XX в. у большинства авторов высокой частотой обладают как *жизнь*, так и *человек*, т.к. для этого среза одной из наиболее крупных маркемных групп является группа фундаментальных понятий (маркемы, связанные с основными, базовыми понятиями: *человек, жизнь, вселенная, время*).

Наибольший интерес представляют маркемы, оказывающиеся высокоупотребительными у представителей всех изучаемых хронологических срезов (см. табл. 4, рис. 4-5).

Таблица 4. Маркемы, высокоупотребительные у представителей всех срезов

Срез	<i>сердце</i>	<i>любовь</i>	<i>смерть</i>	<i>счастье</i>	<i>голос</i>	<i>глаза</i>	<i>солнце</i>	<i>время</i>
XVIII	0,933	0,533	0,466	0,4	0,133	0,066	0,6	0,4
1-XIX	0,866	0,466	0,2	0,666	0,133	0,066	0,066	0,266
2-XIX	0,933	0,466	0,133	0,2	0,4	0,133	0,533	0,2
Срез	<i>сердце</i>	<i>любовь</i>	<i>смерть</i>	<i>счастье</i>	<i>голос</i>	<i>глаза</i>	<i>солнце</i>	<i>время</i>
3-XIX	1	0,466	0,333	0,2	0,266	0,133	0,666	0,266
1-XX	0,9375	0,125	0,1875	0,375	0,3125	0,3125	0,9375	0,125
2-XX	0,733	0,2	0,266	0,066	0,066	0,666	0,533	0,4
3-XX	0,733	0,266	0,2	0,133	0,066	0,6	0,4	0,4

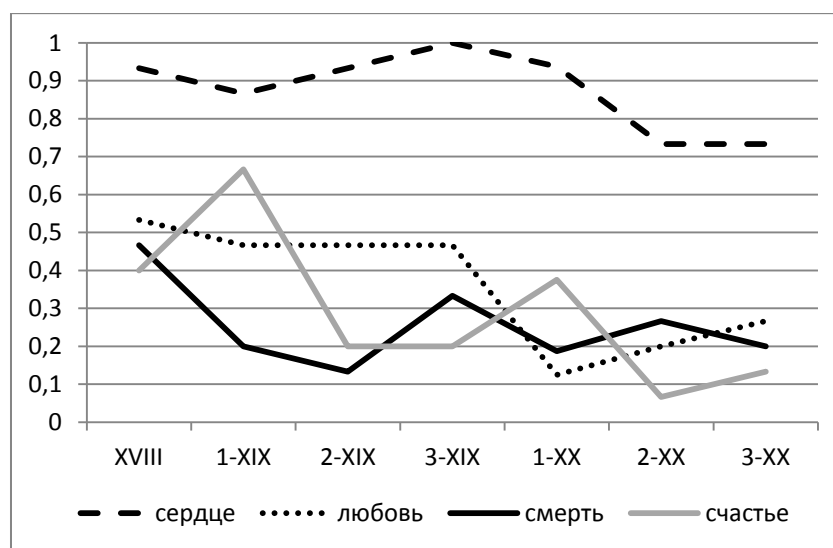


Рис. 4. Маркемы *сердце, любовь, смерть, счастье*: динамика

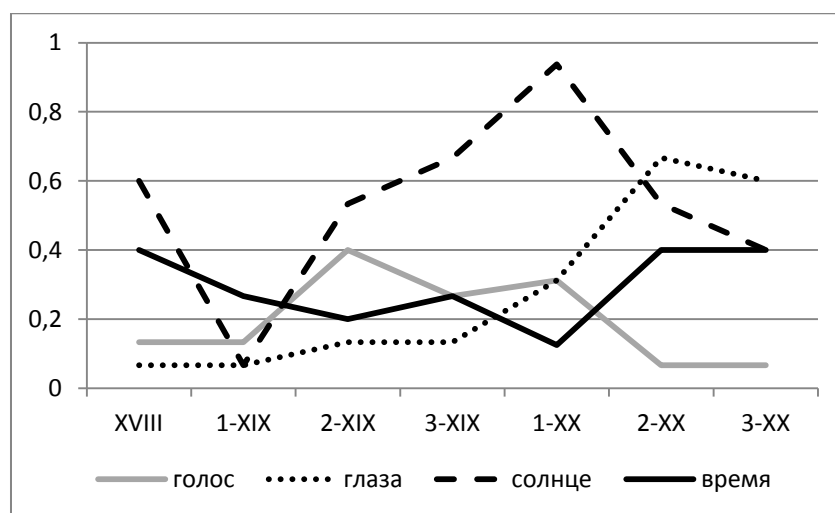


Рис. 5. Маркемы *голос*, *глаза*, *солнце*, *время*: динамика

Маркема *сердце* обладает высокой частотностью у большинства авторов во всех хронологических срезах. Максимум распространённости она достигает в последней трети XIX в., где является одной из наиболее употребительных у всех авторов среза. В целом распространённость этой маркемы колеблется незначительно, чего нельзя сказать о прочих высокочастотных маркемах. Так, маркема *счастье*, встречающаяся в качестве высокочастотной у более чем половины авторов первой трети XIX в., обладает существенной частотой лишь у одного автора в 20-50-х гг. XX в., что может быть связано с трагическими событиями в жизни страны. Маркема *любовь* в качестве высокочастотной типична для XVIII-XIX вв., но в XX в., со снижением доли любовной лирики в общем массиве поэтического творчества, со сменой преобладающей поэтической тематики её распространённость сокращается. Маркема *смерть* как частотная наиболее характерна для XVIII в., что объясняется, возможно, значительной ролью в этом срезе таких жанров, как философская и патриотическая ода. Две другие точки возрастания распространённости маркемы *смерть* приходятся на последнюю треть XIX в., что может быть связано с философской тематикой многих поэтических произведений этого среза, и с общей пессимистичностью, отражавшейся в различных сферах культурной жизни, и на 20-50-е гг. XX в., период войны и репрессий. Маркема *глаза* испытывает резкий скачок распространённости в течение XX в., когда повышается роль группы маркем, связанных с физическим обликом. Число авторов, у которых маркемы *время* и *голос* являются высокочастотными, ни в одном из рассматриваемых хронологических срезов не превышает 6; их распространённость колеблется сравнительно мало. Маркема *солнце*, напротив, испытывает резкие, труднообъяснимые скачки распространённости: в начале XIX в. она является высокоупотребительной лишь у одного автора, а в начале XX в. – у большинства.

Естественно, что маркемы, обычно выступающие в роли высокоупотребительных, отличаются сравнительно малой длиной (5-7 звуков), в большинстве случаев принадлежат к натурфактам,

фундаментальным понятиям, эмоциональным маркемам. Наиболее типичной высокочастотной маркемой является *сердце*.

Полученные нами данные хорошо согласуются с другими результатами исследования системы маркем языка русской поэзии: со сведениями о маркемах, употребляющихся наибольшим количеством авторов, о маркемах с максимальным суммарным ИнТеМ, о маркемах-доминантах хронологических срезов и о наиболее типичных маркемных категориях (Кашкина, 2013: 7-9). Информация о типичных высокоупотребительных маркемах русского поэтического языка вносит свой вклад в исследование языковых универсалий и фреквенталий.

### Литература

1. Кашкина А.В. О классификации маркем // Вестник ВГУ. - Серия Лингвистика и межкультурная коммуникация, 2011. – № 1. – С. 97 - 100.

2. Кашкина А.В. Критерии выделения маркем в русских поэтических текстах//Среди нехоженных путей. Сборник научных статей к юбилею доктора филологических наук, профессора А.А. Кретова / Под ред. И.А. Меркуловой, К.М. Шилихиной. – Воронеж: НАУКА-ЮНИПРЕСС, 2012. – С. 300 - 307.

3. Кашкина А.В. Маркемный анализ языка русской поэзии. – Автореф. ... дисс. канд. филол. наук. – Воронеж, 2013. – 26 с.

4. Кретов А.А. Опыт выявления архетипов поэзии А.В. Кольцова//Лінгвістичні студії. Зб. наук. праць. В. 16./Укл.: Анатолій Загнітко (наук. ред.) та ін. – Донецьк: ДонНУ, 2008. – С. 353 – 366.

5. Кретов А.А. Понятие маркемы: методика выявления и практика использования // Универсалии русской литературы 2. Сборник статей / Под ред. А.А. Фаустова. – Воронеж: НАУКА-ЮНИПРЕСС, 2010. – С. 138 - 153.

6. Кретов А.А., Воронина И.Е., Попова И.В. «Программа выделения тематически маркированной лексики», зарегистрировано в Государственном информационном фонде неопубликованных документов ФГНУ «Центр информационных технологий и систем органов исполнительной власти» (№ 50201000004 от 11.05.2010).

### References

1. Kashkina A.V. On the classification of markems // Vestnik of Voronezh state University. Series of Linguistics and intercultural communication, 2011. - № 1. - p. 97 - 100.

2. Kashkina A.V. Criteria of markems in Russian poetic texts//Among the untrodden ways. Collection of scientific articles for the anniversary of the doctor of Philology, Professor A.A. Kretova / edited by I.A. Merkulova, K.M. Shilikhina. - Voronezh: SCIENCE-UNIPRESS, 2012. - С. 300 - 307.

3. Kashkina A.V. Markem analysis of the language of Russian poetry. – Abstract of PhD thesis in Philology. - Voronezh, 2013. - 26 North

4. A.A. Kretov Experience in identifying archetypes poetry A.V. Koltsov//Лінгвістичні студії. The CG. Sciences. essays. Century

16./UKL.: Анатолій Загнітко (Sciences. amended) and others. - Донецьк: Donna, 2008. - p. 353 - 366.

5. А.А. Kretov Concept маркемы: methods of identification and application /Universals of Russian literature 2. Collection of articles / edited by A. A. Faustov. - Voronezh: SCIENCE-UNIPRESS, 2010. - p. 138 - 153.

6. А.А. Kretov, Voronin I.E., Popov I.V. «selection Program thematically marked vocabulary», registered with the National information Fund of unpublished documents research Center for information technologies and systems of Executive state authorities» (№ 50201000004 from 11.05.2010).

**(0,4 п.л.)**