

УДК 82+81:159.9

НЕКОТОРЫЕ ОСОБЕННОСТИ ПОИСКА СИНЕСТЕТИЧЕСКИХ МЕТАФОР В НАЦИОНАЛЬНОМ КОРПУСЕ РУССКОГО ЯЗЫКА

А.И. Бардовская

Вятский государственный гуманитарный университет, Киров

В статье затрагивается вопрос о ценности электронных корпусов для изучения синестетической метафоры и описываются конкретные шаги при поиске такого рода языковых явлений в Национальном корпусе русского языка.

Ключевые слова: *синестетическая метафора, корпус, Национальный корпус русского языка.*

Синестетическая метафора (далее – СМ), представляющая собой «древнюю, широко распространённую, а, возможно, даже универсальную форму метафоры» [14, с. 292], специфика которой заключается в том, что в ней «слово, значение которого связано с одним органом чувств, употребляется в значении, относящемся к другому органу чувств» [Там же], привлекает сегодня пристальное внимание широкого круга специалистов (об актуальных проблемах современного «синестезиеведения» см., например, [2; 3; 10]). В работах, посвящённых этому языковому явлению, всё большую популярность приобретают сравнительно новые источники информации о языке – электронные корпуса. Это представляется вполне закономерным: результаты ряда исследований (например, [1; 4; 6; 8; 15; 16]) позволяют определить корпус как продуктивный инструмент в изучении СМ.

Однако очевидно, что в эпоху бурно развивающихся компьютерных технологий вновь появляющиеся и совершенствующиеся корпуса неоднородны и нуждаются в сопоставлении и классификации. Возможности для изучения различных языковых явлений, предоставляемые разными корпусами, также нуждаются в оценке. Наконец, отдельного осмысления заслуживают сами принципы, методология работы с корпусами, их отличия от работы с другими источниками сведений о языке.

На текущем этапе нашего исследования мы обратились к такому авторитетному корпусу текстов на русском языке, как Национальный корпус русского языка (далее – НКРЯ) [9]. Не претендуя на решение перечисленных выше задач, попытаемся изложить ряд наблюдений, сделанных нами в ходе работы с НКРЯ. Начнем же с краткой справки о его структуре и составе.

Согласно сведениям на сайте НКРЯ,

«Корпус – это информационно-справочная система, основанная на собрании текстов на некотором языке в электронной форме. Национальный

корпус представляет данный язык на определённом этапе (или этапах) его существования и во всем многообразии жанров, стилей, территориальных и социальных вариантов и т.п.»).

Прежде всего, в НКРЯ включены прозаические оригинальные тексты, представляющие русский литературный язык (с начала XVIII века). В меньшем объеме в нём присутствуют переводные сочинения (параллельно с оригиналом), поэтические тексты и тексты, представляющие разговорную и диалектную формы современного русского языка. Таким образом, наряду с основным корпусом в состав НКРЯ входит ряд подкорпусов (см. табл. 1, представляющую собой фрагмент таблицы с сайта НКРЯ из раздела «Статистика»).

Таблица 1. Число текстов и словоупотреблений в различных подкорпусах НКРЯ

Подкорпус	Число текстов	Число словоупотреблений
Основной	76882	209198275
Газетный	181175	113292003
Диалектный	197	194283
Обучающий	229	664751
Параллельный	370	24022437
Поэтический	41448	6738474
Устный	3034	10122579
Мультимедийный	31741	648576
Всего:	335076	364881378

По мысли создателей корпуса, две отличительные особенности НКРЯ – это представительность, или сбалансированный состав входящих в него текстов, а также наличие так называемой разметки – дополнительной информации об их свойствах. Представляется, что указанные особенности в значительной мере определяют ценность корпуса как инструмента для различных учебных и исследовательских целей и для изучения СМ русского языка, в частности.

Говоря о репрезентативности НКРЯ, следует подчеркнуть, что здесь мы обнаруживаем все типы письменных и устных текстов, причем, по заверению создателей корпуса, представлены они по возможности пропорционально их доле в языке того или иного периода. Обращаясь к НКРЯ, мы получаем доступ к десяткам тысяч художественных текстов разнообразных жанров и нехудожественных текстов различной тематики и сфер функционирования и миллионам словоупотреблений (см. раздел «Статистика» на сайте НКРЯ).

Разнообразная разметка (на данный момент в НКРЯ используется

пять её типов – метатекстовая, морфологическая, синтаксическая, акцентная и семантическая) делает НКРЯ практически непревзойдённым источником сведений о русском языке в ряду аналогичных электронных собраний. Так, весьма репрезентативная электронная библиотека художественных текстов [12], которой мы пользовались в проводившемся нами ранее исследовании СМ [1; 4], не имеет какой-либо разметки. То же самое можно сказать о популярных собраниях текстов, представленных в сети Интернет, например, о «библиотеке Максима Мошкова» [5]. Несколько забегаая вперед, отметим, что особо привлекательным инструментом для изучения СМ НКРЯ становится благодаря частичной семантической разметке. Разные виды разметок свойственны всем общедоступным корпусам русского языка, но сопоставление [11] демонстрирует, что для НКРЯ характерна наиболее полная на сегодняшний день семантическая разметка.

Воспользоваться перечисленными и рядом других достоинств НКРЯ можно в ходе поиска в нём. Представляется, что именно «поиск» является ключевым словом для пользователя электронных корпусов. Наиболее общие рекомендации по грамотному поиску языковых фактов с помощью НКРЯ содержатся на его сайте [9]. С их изучения мы начинали свою работу. Далее, в ходе непосредственной работы в корпусе, специфика интересующего нас языкового явления определяла специфику наших действий. В частности, нами было обнаружено, что поиск СМ в НКРЯ может вестись несколькими, а именно, четырьмя, путями, которые мы попытаемся охарактеризовать в нашей статье.

1. Поиск по *одному компоненту* СМ. Такой вариант поиска СМ в НКРЯ имеет некоторое сходство со сбором языкового материала в ходе работы с традиционными источниками сведений о языке – (бумажными) словарями и текстами. В данном случае, подбирая языковой материал для исследования СМ, исследователь первоначально обращается к разного рода словарям и конструирует лексико-семантическое поле чувственного восприятия в изучаемом языке. Далее, в ходе наблюдений за СМ (как правило, в тексте), выявляются (предпочтительные) направления переносов, делаются заключения об отношениях между различными лексико-семантическими группами (например, говорят о ЛСГ зрения, ЛСГ слуха и т.д.) строится некая типология синестетических переносов (см., например, [7; 13]).

Планируя вести поиск СМ по одному компоненту в НКРЯ, следует иметь готовый список «стимулов», которые, предположительно, будут встречаться в СМ. Введя в поисковую систему корпуса слово, описывающее некое ощущение или восприятие, исследователь получает контексты с ним (их может оказаться не одна тысяча). Контексты анализируются на предмет наличия в них СМ, а далее, в зависимости от целей и задач работы, как и в случае с бумажными носителями, определяются

модели переносов, строятся типологии и т.п.

Подобную работу мы проделывали, изучая СМ с температурным компонентом значения [1; 4], используя для этого электронную библиотеку [12]. Анализируя свой опыт, скажем, что преимущество НКРЯ здесь очевидно: если при работе с электронной библиотекой нам приходилось вводить все словоформы слова по отдельности, то в НКРЯ есть функция лексико-грамматического поиска, дающего контексты со всеми возможными словоформами слова. Но при необходимости можно воспользоваться «поиском точных форм» (например, если вас интересуют только сочетания со словом *холодный* в именительном падеже, единственном числе, мужском роде), в таком случае будут найдены контексты лишь с указанной словоформой.

Если же попытаться вкратце охарактеризовать плюсы и минусы поиска СМ по одному компоненту, то, в первую очередь, отметим, что он весьма трудоёмок (о чём говорят, в частности, цифры в табл. 2). Его результаты в значительной мере зависят от усилий, приложенных исследователем, которые, однако, могут увенчаться более детальной оригинальной классификацией СМ.

Таблица 2. Сочетания со словом *холодный* при разных вариантах поиска СМ в НКРЯ

	Количество документов	Количество вхождений	Количество СМ (первые десять контекстов)
Без указания грам. и сем. признаков	7788	37419	2
С указанием грам. и сем. признаков	157	176	9

2. Поиск определённого *целого синестетического сочетания* интересен, в первую очередь, для выяснения статистических данных относительно употребления той или иной конкретной СМ. С другой стороны, работая с НКРЯ таким образом, мы получаем возможность детализации контекстов, в которых функционирует изучаемая СМ.

Какие же возможности для ведения статистики поиска предлагает НКРЯ? В первую очередь, к результатам любого поиска предлагается «статистика по метаатрибутам» и также «распределение по годам». Статистика по метаатрибутам представляется в виде таблиц и учитывает различные характеристики текстов, в которых была обнаружена СМ: информация об авторе и поле автора, сфера функционирования, тип, те-

матика и жанр текста. Распределение по годам представляется в виде графика, демонстрирующего частоту появления того или иного языкового явления на миллион словоформ.

По умолчанию поиск в НКРЯ ведётся в основном подкорпусе. Однако НКРЯ предлагает своим пользователям создавать подкорпусы текстов, в которых они будут вести работу. Например, для нашего исследования представляет интерес сопоставить тексты, написанные на рубеже прошлого и позапрошлого веков (когда СМ выделилась в сферу отдельного научного интереса и наблюдалось её особо активное использование) и современные тексты (тексты последних 10–20 лет, характеризующихся как «синестетический бум»); СМ в текстах, написанных женщинами и мужчинами; СМ в текстах основного и поэтического корпусов; понаблюдать за СМ в параллельном корпусе. Эти и многие другие возможности предоставляет НКРЯ своим пользователям, составляющим подкорпусы текстов, пригодные для различных исследований.

Возвращаясь к характеристикам поиска определённого целого синестетического сочетания, отметим, что технически такой поиск осуществляется так же, как поиск по одному компоненту СМ, с той лишь разницей, что вводить в строку поиска здесь следует всё словосочетание (например, *холодный взгляд, тёплый цвет*). И поиск точных форм, и лексико-грамматический поиск позволяют находить СМ любой протяжённости (т.е., пользуясь НКРЯ, в данном случае можно не ограничиваться традиционными двухкомпонентными СМ; другое дело, что трёх- и более компонентные СМ повторяются крайне редко). Пользуясь лексико-грамматическим поиском, можно также задать расстояние между словами в СМ.

3. Поиск по одному компоненту СМ с использованием *разметки*. Разметка в НКРЯ весьма полезна при поиске по одному компоненту СМ, поскольку позволяет ограничить параметры поиска. В данном случае осуществляется лексико-грамматический поиск. В строке «Слово 1» (или «Слово 2», в зависимости от того, по какому компоненту СМ будет вестись поиск) вводится слово-стимул. Для второго слова задаются морфологические и семантические параметры.

К примеру, мы хотим получить СМ со словом *холодный*. Вводим это слово в строку «Слово 1» (ведь известно, что для русского языка типичны СМ модели «прилагательное+существительное», а не наоборот). «Слово 2» нам неизвестно, однако, мы можем указать его параметры, отметив галочкой «существительное» в строке «Грамм. признаки» (в принципе, для поиска СМ этого достаточно, однако, если есть задача отобрать сочетания «прилагательное+существительное» с определёнными грамматическими параметрами, то можно их уточнить, указав падеж, число, род и т. д.) и «форма», «цвет», «вкус», «запах» в строке «Семант. признаки» (результаты поиска см. выше в табл. 2).

В зависимости от целей и задач работы, можно задать не все используемые в НКРЯ семантические признаки, а один или два. В таком случае могут оказаться необходимыми некоторые корректировки грамматических признаков. К примеру, исследуя СМ модели ТЕМПЕРАТУРА → ЦВЕТ, в строке «Грамм. признаки» следует отмечать не только существительное, но и прилагательное, поскольку здесь мы можем ожидать появление сочетаний типа *холодный синий* («прилагательное+прилагательное»).

Этот способ поиска является более «энергосберегающим» для исследователя. Заданные параметры существенно сокращают круг выбираемых поисковой системой контекстов. С другой стороны, не следует забывать, что семантическая разметка в НКРЯ лишь частична и в получаемых таким образом контекстах СМ реализуется не всегда.

4. Поиск СМ с использованием разметки предполагает указание лишь грамматических и семантических признаков слов, участвующих в синестетическом сочетании. Рассмотрим его особенности, опять же, на примере типичной для СМ модели «прилагательное+существительное».

В строках «Грамм. признаки» в данном случае отмечаем «прилагательное» и «существительное». В строках «Семант. признаки» – «форма», «цвет», «вкус», «запах», «температура» для прилагательного и «звук», «цвет», «свет», «запах и вкус» и «температура» для существительного. Получаем список контекстов, состоящий из 9441 документа, 52812 вхождений. Статистика по полученному списку в целом соответствует положениям, выдвигавшимся исследователями СМ. Так, согласно графику распределения таких сочетаний по годам, пик их употребления приходится на начало прошлого века, явный же спад начинается в середине прошлого века. Место «чемпиона» по использованию СМ занимает с большим отрывом от других М. Горький (57 документов, 465 словоформ). Более типичны СМ для авторов-мужчин (3439 документов, 11519 словоформ; ср. с жен.: 606 документов, 1561 словоформ), для художественной литературы (2381 документов, 9763 словоформы).

Однако, обращаясь к анализу сочетаний, отобранных компьютером, обнаруживаем, что многие из них не относятся к числу СМ. Так, наряду с такими «классическими» СМ, как *тонкие ароматы, чистые голоса*, мы видим сочетания *коричневый цвет, чёрный цвет*, представляющие не что иное, как использование обоих входящих в них слов в их прямом значении (на что указывает и контекст их употребления). А это ещё раз показывает, что внимание исследователя – одно из главных условий продуктивной работы с корпусом.

Список литературы

1. Бардовская А.И. Средства номинации синестетических соощущений (на материале английских и русских художественных текстов) [Текст] : автореф. дис. ... канд. филол. наук : 10.02.19 / А.И. Бардовская ; Тверской государственный университет. – Тверь, 2005. – 20 с. – На правах рукоп.
2. Бардовская А.И. Современные тенденции изучения синестетической метафоры [Текст] / А.И. Бардовская // Вестник Тверского государственного университета. – Серия «Филология». – 2010. – № 15. – Вып. 4 «Лингвистика и межкультурная коммуникация». – С. 168–181.
3. Бардовская А.И. Проблема синестезии и интеграционные тенденции в языковедческих исследованиях [Текст] / А.И. Бардовская // Вестник Вятского государственного гуманитарного университета. – 2011. – № 1(1). – С. 131–137.
4. Бардовская А.И. Синестетическая метафора с компонентом значения «температура» в русском и английском языках: опыт работы с электронными библиотеками [Текст] / А.И. Бардовская // Вестник Вятского государственного гуманитарного университета. Филология и искусствоведение. – 2011. – № 2 (2). – С. 78–85.
5. Библиотека Максима Мошкова [Электронный ресурс] : личный сайт. – Электрон. дан. – [М.], 1994 / URL: <http://lib.ru/> – Дата обращения: 2012 г. – Загл. с экрана.
6. Колодкина Е.Н. Вкусовая синестетическая метафора [Электронный ресурс] / Е.Н. Колодкина / URL: www.gramota.net/materials/2/2009/1/25.html. – Загл. с экрана.
7. Мерзлякова А.Х. Вторичные функции перцептивных прилагательных (на материале английского, русского и французского языков) [Текст] / А.Х. Мерзлякова. – Ижевск : Изд. дом «Удмуртский университет», 2003. – 198 с.
8. Молодкина Ю.Н. Синестетическая метафора запаха (корпусное исследование) [Текст] : автореф. дис. ... канд. филол. наук : 10.02.19 / Ю.Н. Молодкина ; Курский государственный университет. – Курск, 2010. – 20 с. – На правах рукоп.
9. Национальный корпус русского языка [Электронный ресурс] : личный сайт. – Электрон. дан. – [М.], 2009 / URL: <http://www.ruscorg.ru/> – Дата обращения: 2012 г. – Загл. с экрана.
10. Прокофьева Л.П. Синестезия в современной научной парадигме [Текст] / Л.П. Прокофьева // Известия Саратовского университета. – 2010. – Т. 10. – Серия «Филология. Журналистика». – Вып. 1. – С. 3–10.
11. Резникова Т.И., Копотев М.В. Лингвистически аннотированные корпуса русского языка (обзор общедоступных ресурсов) [Электронный ресурс] / Т.И. Резникова, М.В. Копотев / URL:

- <http://ruscorpora.ru/sbornik2005/04reznikova.pdf>. – Загл. с экрана.
12. Русская литература: от Нестора до Маяковского [Электронный ресурс]. – Электрон. дан. – М. : ДиректМедиа, 2003. – CD-ROM. – Загл. с контейнера.
 13. Степанян Т.Р. Синестезические метафоры русского языка (прилагательные чувственного восприятия) [Текст] : дис. ... канд. филол. наук : 10.02.01 / Т.Р. Степанян ; Институт русского языка АН СССР. – М., 1997. – 184 с.
 14. Ульманн С. Семантические универсалии [Текст] / С. Ульманн // Новое в зарубежной лингвистике. – М. : Прогресс, 1970. – Вып. V. – С. 250–299.
 15. Day, S. Synaesthesia and Synaesthetic Metaphors [Электронный ресурс] / S. Day / URL: <http://psyche.cs.monash.edu.au/v2/psyche-2-32-day.html>. – Загл. с экрана.
 16. Salzinger, J. The Sweet Smell of Red – An Interplay of Synaesthesia and Metaphor in Language [Электронный ресурс] / J. Salzinger / URL: www.metaphorik.de/18/salzinger.pdf. – Загл. с экрана.

SYNAESTHETIC METAPHOR: SOME PECULIARITIES OF SEARCH THE RUSSIAN NATIONAL CORPUS

A.I. Bardovskaya

Vyatka State University of Humanities, Kirov

The article stresses the idea of the importance of using the language corpora in the synaesthetic metaphor studies and describes the steps taken while searching the Russian National Corpus.

Keywords: *synaesthetic metaphor, corpora, Russian National Corpus.*

Об авторе:

БАРДОВСКАЯ Анастасия Игоревна – кандидат филологических наук, доцент, докторант кафедры английского языка Тверского государственного университета, e-mail: nastya978@inbox.ru