

Г.И. Кустова

**ЭЛЕКТРОННЫЙ СЛОВАРЬ
СТЕПЕННОЙ СОЧЕТАЕМОСТИ НА БАЗЕ
НАЦИОНАЛЬНОГО КОРПУСА РУССКОГО ЯЗЫКА¹**

Национальный корпус русского языка можно использовать не только в «онлайновом» режиме (как источник языковых примеров, средство проверки гипотез и т.п.), но и для создания новых Интернет-ресурсов и электронных продуктов, например, его можно использовать как основу для создания различных электронных словарей (баз данных). В настоящее время на базе Корпуса создается целая серия словарей – частотный словарь, словарь служебных глаголов, словарь редких слов. Одним из таких словарей, о котором и пойдет речь в данной работе, является словарь количественно-степенных сочетаний, в которых адverbиальный или адъективный определитель (модификатор) выражает значение высокой степени, ср. *глубокое удовлетворение, вопиющее безобразие, чертовски обаятельный, несказанно рад* и под. Далее для краткости мы будем называть соответствующие модификаторы степенными словами, а создаваемый электронный словарь – Словарем степенной сочетаемости.

Значение такого Интернет-ресурса очевидно. Прежде всего, электронный словарь степенной сочетаемости, как и любой другой словарь, может использоваться носителями языка и иностранцами в качестве справочного источника.

Кроме того, Словарь может служить важным подспорьем для профессиональных лингвистических исследований. В центре внимания лингвистов традиционно оказывалась степенная сочетаемость качественных прилагательных и наречий (*крайне печально, чрезвычайно интересный*). Однако степенные слова могут

¹ Работа выполняется при поддержке РГНФ, проект № 08-04-00183а.

сочетаться также с существительными и глаголами разных семантических классов, и по поводу таких сочетаний есть множество не только не решенных, но даже еще не поставленных вопросов. Словарь степенной сочетаемости может служить полигоном для постановки и решения таких вопросов.

С другой стороны, словарь имеет и совершенно конкретное прикладное назначение. Он создается в рамках проекта Министерства образования РФ и будет использоваться в сфере обучения русскому языку как родному и как неродному. Степенная сочетаемость вызывает определенные затруднения даже у взрослых носителей языка (об этом свидетельствуют получившие широкое распространение в устной речи и даже в языке СМИ обороты вроде **очень прекрасно*, которые, наряду с неправильным употреблением компаратива **более лучше*, стали характерной приметой современного языка). В еще большей степени это относится к школьникам и студентам.

Особая трудность освоения степенной сочетаемости состоит в том, что она в значительной мере лежит в сфере идиоматики. Лишь очень небольшое количество «основных» степенных слов – *очень, ужасно, страшно, сильно, крайне* – имеют относительно свободную сочетаемость (*ужасно добрый, страшно устал, сильно простудился, крайне недоволен*), хотя и такие общераспространенные и частотные показатели значения высокой степени имеют определенные сочетаемостные ограничения (ср. **сильные трудности – большие, колоссальные трудности, *крайний скандал – страшный, жуткий, грандиозный скандал, ?страшное недоумение – крайнее недоумение*). Большинство же степенных слов имеет идиоматичную сочетаемость. Идиоматичные сочетания составляют трудность не только для иностранцев, но и для носителей языка, особенно для детей и подростков, еще не обладающих развитыми языковыми навыками и достаточной культурой речи. Школьники часто не знают, «как правильно сказать», и допускают многочисленные «идиоматические ошибки» (особенно в пись-

менных текстах – сочинениях, изложениях), например: *безоговорочное решение (ср.: безоговорочное согласие, бесповоротное решение); *оголтелая ложь (ср.: беспардонная ложь, оголтелая травля), *беспардонное пьянство (ср.: беспробудное пьянство, беспардонная ложь).

Среди степенных сочетаний есть «абсолютные» идиомы вроде *закадычный друг, истошный крик, гробовая тишина, наотрез отказаться, плакать навзрыд, дура набитая, мертвенно бледный, мертвецки пьян, насущно необходимый* и под. и есть такие, в которых степенные слова допускают более широкий, но тем не менее ограниченный круг контекстов. Например, носители языка (а также изучающие язык) должны знать (запоминать), что *боль, досада, жажда, жалость, зависть* бывает и *жгучая* и *острая*, *нехватка* и *потребность* – *острая*, но не *жгучая*, а *мороз* и *брюнет* – *жгучий*, но не *острый*; что можно сказать *глубокое / крайнее / бурное возмущение* и *вопяющее / грубое / злостное нарушение*, но нельзя **глубокое / крайнее / бурное нарушение* и **вопяющее / грубое / злостное возмущение*. При этом определенная семантическая мотивированность в идиоматических сочетаниях есть, т.е. вполне объяснимо, что для выражения значения высокой степени используются прилагательные *острый, резкий, жгучий, горячий, страшный, ужасный, невыносимый* и под.¹ Однако не удастся в достаточно простых правилах сформулировать и объяснить, почему по-русски говорят, например, *неоценимый вклад*, но не говорят *?неоценимые достижения* (при том что *ценные достижения* в Корпусе представлены).

Для учащихся идиоматические сочетания представляют особые трудности еще и потому, что многие из этих сочетаний имеют книжный характер (*неуклонный рост, непререкаемый автори-*

¹ О семантических источниках и закономерностях формирования значений высокой степени см. в работе: *Кустова Г.И.* Количественные значения качественных слов // Логический анализ языка. Квантификативный аспект языка. М., 2005. С. 295–305.

тет, кровная заинтересованность), а книжная, официальная речь усваивается труднее и позже, чем разговорная.

Развитие и совершенствование языковых навыков учащихся в сфере идиоматики предполагает использование, наряду с традиционными бумажными словарями, современных информационных ресурсов, базирующихся на всем богатстве данных Национального корпуса русского языка и позволяющих проводить выбор языковых средств, их сравнительный анализ и верификацию экспериментально построенных сочетаний по Корпусу в онлайн-режиме.

Лингвистической основой для создания словаря степенной сочетаемости на базе Корпуса является имеющаяся в Корпусе морфологическая и семантическая разметка.

Создание словаря включает следующие технологические этапы.

1) В семантической разметке Корпуса имеется помета «степень», соответственно, в Корпусе уже есть слова с данной пометой (*очень, колоссальный, крайний, беззаветно, безгранично* и т.п.). Для целей создания словаря список степенных слов был расширен (ср. *безнадежно (устаревший), мертвенно (бледный)* и др.) и в итоге составил около 700 единиц. При пополнении списка был использован «Словарь усилительных словосочетаний...»¹, а также различные толковые словари и словари сочетаемости.

2) Из корпуса извлекаются двусловные сочетания со степенными словами.

Поскольку многие степенные слова многозначны, т.е. имеют и другие, не степенные, значения, в список двусловных сочетаний наряду с собственно степенными (например, *глубокое отчаяние, крайнее недоумение*) попадут и другие (*глубокий колодец, крайнее окно*).

¹ Сущинский И.И. Словарь усилительных словосочетаний русского и немецкого языков. М., 1997.

3) Из общего списка экспертом вручную выбираются собственно степенные сочетания.

Поскольку словарь существует в формате электронной таблицы, а извлеченные из Корпуса сочетания снабжены морфологической (равно как и семантической, см. ниже) разметкой, то пользователь может получить не только общий алфавитный список степенных сочетаний, но и отдельные списки по всем возможным моделям степенных сочетаний:

а) наречие + прилагательное (*крайне непоследовательный, глубоко несчастный, мертвецки пьян*);

б) наречие + наречие (*необычайно просто, поразительно быстро*);

в) наречие + предикатив (*глубоко безразлично, мучительно стыдно, невообразимо тесно, невыносимо тяжело, нестерпимо больно*);

г) наречие + глагол (*крепко запомнить, наотрез отказаться, насквозь промокнуть, разительно отличаться*);

д) прилагательное + существительное (*закоренелый холостяк, заядлый спорщик, кристальная честность, неземное блаженство, ярый приверженец*).

Кроме того, о конкретных единицах, входящих в словарь, пользователь сможет получить информацию двух типов: (1) с какими словами сочетается данный степенной показатель (ср. **круглый**: *двоечник, дурак, идиот, невежда, неуч, отличник, сирота, круглая*: *бездарность, дура*); (2) какие степенные показатели возможны при данном слове (например: **честность**: *абсолютная, большая, исключительная, кристальная, невероятная, невероятная, необычайная, непогрешимая, огромная, стопроцентная*).

Таким образом, технологические этапы создания словаря более или менее ясны. Однако для словарей такого типа принципиально важным является содержательное наполнение, формирование словника, которое, в свою очередь, основывается на определенных теоретических представлениях о том, какие слова выражают степень и что считать степенными значениями.

* * *

К словам с «усилительным» значением (см., например, «Словарь усилительных словосочетаний...»¹) относят не только собственно степенные слова (*очень*), но и слова со значением большого размера (*огромный, гигантский, колоссальный*), предела (*абсолютно, совершенно*), полного охвата множества (*поголовно, повальный*). Применимость этих определителей к определяемым словам (которые мы для краткости будем называть контекстами) зависит от семантики самих определяемых слов.

Качественные прилагательные и наречия могут принимать показатели степени по определению (прилагательное (наречие) называется качественным, если оно сочетается с показателями степени). Хотя сочетаемость таких слов, в принципе, семантическая, она не является совершенно однородной: разное семантическое устройство предполагает и разный тип сочетаемости – собственно степенную (*очень длинный*) либо предельную (*совершенно сухой*) (о значениях степени и предела см. в работе Ю.Д. Апресяна²). Есть различия и внутри степенной сочетаемости, ср.: *глубоко безразлично* – **глубоко неприлично*, **ужасно безразлично* – *ужасно неприлично*.

Гораздо более разнообразной и, с другой стороны, менее изученной является степенная сочетаемость глаголов и существи-

¹ *Сущинский И.И.* Словарь усилительных словосочетаний...

² *Апресян Ю.Д.* Лексическая семантика: Синонимические средства языка. М., 1974.

тельных. Здесь возникает множество конкретных (по отдельным словам) и общих (по целым классам слов) вопросов, связанных, во-первых, с тем, какие абстрактные объекты и ситуации допускают количественно-степенные характеристики и, во-вторых, с тем, какие слова с исходно качественной семантикой предрасположены к развитию степенных значений и в каких контекстах такие значения формируются.

Существительные, принимающие показатели высокой степени, весьма разнообразны по семантике. Они могут быть обозначениями лица (*отъявленный негодяй, махровый реакционер, прожженный мошенник*), эмоциональных состояний и реакций (*райское блаженство, глубокое отвращение, неописуемый ужас, жгучий стыд*), свойств, способностей и проявлений человека (*невероятное высокомерие, звериная жестокость, ангельское терпение*), состояний и явлений природы (*адская жара, лютый мороз*), они могут иметь посессивную семантику (*баснословная прибыль, сказочное богатство*), модальную (*крайняя нужда, насущная необходимость, настоятельная потребность*), оценочную (*чудовищная безвкусица, невиданное великолешие*) и др.

Глаголы, являющиеся контекстами степенных слов, также относятся к разным семантическим классам: ментальные (*безгранично доверять, начисто забыть, крепко запомнить, слепо верить*), эмоциональные (*панически бояться, глубоко возмущать, беззаветно любить, смертельно обидеться*), посессивные (*сказочно разбогатеть*), речевые (*беззастенчиво врать*) и др.

Прежде всего, контексты, в зависимости от семантики, требуют разных типов количественно-степенных определителей. Некоторые слова (и, соответственно, обозначаемые ситуации или абстрактные объекты) характеризуются с точки зрения степени (*адская жара; смертельная обида; крайняя отсталость; разительно отличаться; безграничное терпение, ср. *полное терпение*); некоторые – с точки зрения величины (*огромный авторитет; колоссальные достижения*), некоторые требуют модифика-

торов с предельным значением (*абсолютно игнорировать, полностью доверять, вконец испортить, полная невиновность, абсолютная непригодность*, ср. **абсолютная жара*); некоторые – со значением полного охвата множества (*поголовная грамотность, повальное увлечение*). При этом иногда одни и те же ситуации могут характеризоваться и с точки зрения «величины» (*большое, огромное разочарование*), и с точки зрения степени (*глубокое, сильное, жестокое, острое разочарование*), и с точки зрения предела, исчерпанности (*полное разочарование*).

Кроме того, слова разных семантических групп имеют свои предпочтения в сфере степенной сочетаемости (ср.: *адская жара, смертельная обида, катастрофический спад*, но не **смертельный спад, *адская обида, *катастрофическая жара*), хотя есть степенные слова, которые сочетаются с разными семантическими классами контекстов (*необычайная активность, жара, красота, щедрость, яркость, необычайное волнение, благородство, необычайный успех, шум; страшный голод, грохот, дефицит, мороз, переполох, пожар, скандал, страшное волнение, искушение, поражение, потрясение, страшная путаница, резня, слабость, суматоха, теснота, тоска, тревога*).

Здесь возникает множество вопросов, связанных с семантической избирательностью, широкой и узкой сочетаемостью, связью типа модификатора и типа контекста. В решении этих вопросов, несомненно, может быть использован Словарь степенной сочетаемости.

С другой стороны, сами количественно-степенные слова также образуют весьма неоднородную группу и могут рассматриваться с разных точек зрения. Во-первых, у них есть разные подклассы по типу количественной характеристики, о которых мы уже упоминали выше (степень, предел, величина, полный охват множества). Во-вторых, есть проблема разграничения количественных и качественных значений.

Некоторые слова утрачивают исходную сочетаемость, «покидают» свой класс сочетаемости, применяются к «чужим» объектам и ситуациям, и это является сигналом того, что они из качественных характеристик превратились в степенные: *крайний дом – крайнее раздражение; бешеный бык – бешеная скорость*. Но наряду с подобными словами есть и другие, которые имеют исходную, «собственную» сочетаемость, т.е. являются качественными характеристиками с совершенно прозрачной семантической мотивированностью, но при этом имеют и значение высокой степени, поскольку такие качественные характеристики присоединяются к ситуации только в том случае, если она реализуется в «интенсивном» или даже «максимальном» варианте: *беззащитное вранье, безжалостная ирония*. В таких случаях не всегда просто решить, какие сочетания включать в Словарь, а какие – нет, ср. *панический страх* и *паническое бегство*, *безнадежно отстал* и *безнадежно болен*. Другой случай – характеристика указывает на следствие, но следствие из высокой степени, поэтому она сама становится показателем высокой степени: *неузнаваемо изменился* – так сильно, что нельзя узнать; *безутешное горе* – такое сильное, что нельзя утешить; *непримиримая вражда* – такая сильная, что невозможно примирение. Наконец, степенные слова весьма разнообразны по происхождению и по механизмам образования количественно-степенных значений у исходно не степенных слов. Степенные значения могут развиваться у слов с параметрической семантикой (*высокий, глубокий*), с оценочной семантикой (*ужасный, страшный, чудовищный, райский, адский*; при этом значение оценки также может быть не исходным, не первым), с экспериенциальной¹ семантикой (*неслыханный, невиданный, невообразимый, невыносимый*) и др.

¹ Кустова Г.И. Типы производных значений и механизмы языкового расширения. М., 2004.

После специального исследования список степенных слов будет снабжен дополнительными пометами, соответствующими частным семантическим классам и группам внутри общего корпуса степенных слов. Не исключено, что будет несколько помет – по разным параметрам (линиям анализа).

Между всеми этими аспектами и параметрами могут существовать связи, которые необходимо обнаружить и объяснить. И это является одной из главных научных целей создания Словаря степенной сочетаемости.

Лингвистически обоснованное выделение классов и подклассов степенных слов, а также наличие заимствованной из Корпуса морфологической и семантической разметки у их контекстов дает большие возможности пользователю, позволяет получать самую разную лингвистически релевантную информацию, например:

– в каких контекстах (зафиксированных в корпусе) некоторое слово имеет значение степени (ср. *безумные цены* vs. *безумные мечты*, *глубоко раскаиваться* vs. *глубоко копать* и т.д.);

– какие степенные прилагательные не имеют соответствующих наречий (*неуемная жадность* – **неуемно жадный*, *отпетый лентяй* – **отпето лентяйничать*); разумеется, есть и степенные наречия, не имеющие соответствующих прилагательных, но это объясняется в первую очередь словообразовательной моделью: *напрочь*, *наотрез*, *навзрыд* и т.п.; кроме того, бывают расхождения в степенной сочетаемости у однокоренных слов: *неузнаваемо изменился* – ?*неузнаваемые изменения*; *нестерпимый голод* – **нестерпимо голодать*;

– с какими частями речи сочетается степенной показатель, например: с прилагательными, наречиями и глаголами: *очень*, *чрезвычайно*, *крайне* (*важный*, *важно*, *устать*); с прилагательными и наречиями: *невообразимо*, *потрясающе*, *поразительно* (*быстрый*, *быстро*); разумеется, подобные степенные слова имеют и адъективный вариант – *невообразимый*, *потрясающий*, *по-*

разительный, – однако однокоренные прилагательные и наречия считаются в Словаре разными единицами); с глаголами: *истошно кричать, круто измениться, люто ненавидеть*;

– какие семантические классы глаголов / существительных допускают степенные слова (этот вопрос релевантен именно для глаголов и существительных, поскольку для прилагательных и наречий достаточно качественности);

– с какими степенными модификаторами могут сочетаться слова определенного семантического класса, например, слова со значением лица (*заядлый болельщик, грязный развратник, отъявленный плут*), эмоциональной реакции и отношения (*смертельно обидеться, страстно любить*), оценки (*полная ерунда, страшная гадость*), с ментальным значением (*слепо верить, напрочь забыть*), с посессивным значением (*фантастически разбогатеть*) и т.д.

Подведем итоги. Электронный Словарь степенной сочетаемости, создаваемый на базе Национального корпуса русского языка, ориентирован на самые разные группы пользователей – от школьников и иностранцев, изучающих русский язык, до профессиональных лингвистов. Кроме информации общего характера, из словаря можно будет получить более специальную информацию, представляющую интерес для лингвистических исследований.