

## КОРПУСНАЯ VS. ПСИХОЛИНГВИСТИКА: КОНФРОНТАЦИЯ ИЛИ КООПЕРАЦИЯ

О. О. Борискина

*Воронежский государственный университет*

Поступила в редакцию 20 июня 2009 г.

**Аннотация:** *статья затрагивает вопросы выбора исследовательской стратегии и тактики для описания языкового сознания. Рассматривается возможность проверки корпусных данных экспериментальными методами. Предложены результаты изучения распределенности непередметных имен в одном из криптоклассов английского языка.*

**Ключевые слова:** *психолингвистика, эксперимент, корпусная лингвистика, языковая картина мира, криптокласс имен.*

**Abstract:** *the paper deals with the choice of research strategy and tactics for description of «what language categories reveal about the mind». The perspective of corpus-based and experimental methods intersection is discussed. Distribution of abstract nouns in cryptotype «Liquidus» is described.*

**Key words:** *psycholinguistics, experiment, corpus linguistics, linguistic categorization, noun cryptotype.*

### Корпус или эксперимент

Настоящая статья появилась как результат рефлексии на обсуждение моего доклада на XVI Международном симпозиуме по психолингвистике и теории коммуникации (Москва, 15—17 июня 2009 г.). Бесспорно, порождение текста по правилам языка, в норме контролируемое языковой интуицией говорящего, является предметом изучения как в психолингвистике, когнитивной и теоретической лингвистике, так и в компьютерной лингвистике. На пересечении интересов данных направлений научного поиска лежит и предмет данного исследования — скрытая словарная категория (*криптокласс имен*). Метод криптоклассного распределения имен интегрирует качественный и количественный анализ текстового материала, полученный из корпусов BNC (*British National Corpus*) и COCA (*Corpus of Contemporary American English*). Наши попытки проникнуть в языковое сознание англоговорящих, опираясь на тексты, хотя и оторванные от коммуникативной ситуации, но свидетельствующие об автоматизированном навыке говорящих сочетать слова, позволяют выстраивать модели криптоклассного хранения знания, конструировать гипотезы и прогнозы о системных и функциональных особенностях исследуемых явлений. О возможности проверки таких гипотез и прогнозов экспериментальными методами говорили многие участники симпозиума: «...любой автоматизированный навык, особенно если он относится к функционированию единиц высших уровней языковой системы ... может

быть поставлен в экспериментальную ситуацию опознания» [1, с. 43].

На симпозиуме поднимались вопросы востребованности и эффективности экспериментальных методов анализа [см., например, 1, 2]. Заметим, что в независимости от симпозиума на желательность проверки результатов исследований корпусов и интернет-коммуникации экспериментальными методами неоднократно указывал Ю. Д. Апресян. Ср. также убеждение А. А. Леонтьева в том, что эксперимент служит для поиска и проверки теоретических знаний [3].

Действительно, «как показывает опыт экспериментальной работы, в языковом сознании правила (в частности, синтаксические) являются более лабильными и вместе в тем привязанными к определенным типам контекста, они ... неразрывно связаны со знаниями о мире и в целом с когнитивным и речедетельностным опытом человека, но не менее значима их пронизанность эмоциональными, субъектно-апперцепционными установками языковой личности» [1, с. 45]. Именно такими установками личности можно объяснить ее способность к словотворчеству, окказиональному словосочетанию, что находит отражение в корпусах и интернет-коммуникации. Такой материал непозволительно игнорировать, но следует четко определить его ценность (наряду с узусом и нормой) и релевантность поставленной задаче.

Дуализм языкового сознания представлен в отечественной психолингвистике различием картины мира и образа мира. Первое связывают с коллективным языковым сознанием (ЯКМ как факт культуры), а второе — с индивидуальным (лишь отчасти совпадающим с коллективным) опытом познания мира

через приобщение к культуре (образ мира как достояние субъекта, которое может принципиально отличаться от индивида к индивиду) [4, с. 82; 5, с. 297]. Для выявления степени такого совпадения и уникальности образов мира отдельных испытуемых используются психолингвистические эксперименты разных видов. Но разве корпус (равно как и интернет) не является коллекцией текстов, авторы которых в вербальной форме отражают уникальность своих образов мира?! Непредвзято оценить значимость языкового материала, представленного в самом полном виде в ресурсах компьютерного формата (БД, корпусах, Web-2), поможет нам четырехчастная схема отношений между знаком и объектом реальной или сконструированной в культуре действительности, предложенная Г. П. Мельниковым: «Знак — образ знака —\* образ денотата — денотат». «Языковой знак условен, но необходимое звено в означивании — безусловная, неконвенциональная, окказиональная ассоциация по сходству, возникшая в момент опознания знака» (звездочкой помечена окказиональная ассоциация по сходству. — Б. О.) [6, с. 242].

На наш взгляд, противопоставление корпусного и экспериментального подходов к изучению языкового сознания ничего не может дать науке. Только объединение этих подходов сможет повысить обоснованность и вескость аргументов теоретического исследования. Рассмотрим сначала роль корпусных технологий в опознании и описании коллективного (т.е. закрепленного в культуре, узусе) образа языкового знака, который находится в ассоциации по сходству с образом денотата, получаемого в результате гештальтного опознания денотата субъектом. А затем определим роль лингвистического эксперимента, основанного на наложении образа мира индивида на картину мира языкового коллектива.

### Всё течет, всё изменяется!?

#### Так ли это по данным корпусов?

Если опираться на философские воззрения (например, Гераклита) или мифы народов мира, то течение — одна из форм существования материи. Если под течением (перемещением по образу и подобию жидкой субстанции) понимать один из способов изменения, то очевидно в **языковой картине мира** можно обнаружить свидетельства «текучести» VS. «нетекучести» сущностей. При этом речь идет не о реальном положении дел, а о наивном, мифологизированном, неправдоподобном, аморфном и в силу этого ускользающем от пытливого исследователя представлении о некоей социальной абстракции, типа независимость *independence*, справедливость *justice* или обсуждение *negotiation*.

Что таит мифологизированное сознание англоязычного социума?

Разрабатываемый в рамках проекта кафедры теоретической и прикладной лингвистики\*, метод криптоклассной распределенности именной лексики позволяет через описание ядра и периферии криптокласса английского языка «Liquidus» выявить культурные реалии, которые тысячелетиями уподоблялись жидкому, текучему (Liquidus) в языковом сознании англофонов.

Ниже списком мы приводим имена, которые по данным корпусов не моделируются в англоязычном сознании по образу и подобию водной стихии: *arrest, aspect, basis, campaign, cause, challenge, community, crime, deficit, department, domination, east, election, environment, example, facility, factor, failure, finance, focus, front, government, incident, independence, inflation, injury, instance, insurance, kind, law, length, management, marriage, means, ministry, mission, negotiation, north, number, option, pair, part, piece, possibility, programme, progress, purpose, rate, resolution, rival, science, side, solution, south, state, sport, technique, test, treatment, trust, type, union, west, youth*.

Неожиданно, на первый взгляд, прототипом (т.е. ядерным представителем криптокласса), наряду с водой (*water*) и кровью (*blood*), оказалось имя *life* (жизнь). Однако более пристальное изучение вопроса проясняет ситуацию. Ничего удивительного ни с точки зрения научного знания, ни с позиций мифологизированного сознания здесь нет. Жизнь зародилась и долгое время **протекала** в воде, а в древнефиванской теологии, например, сотворение земного диска из первобытного океана Нун отождествлялось с жизненным дыханием.

Попытаемся интерпретировать результаты качественного и количественного анализа корпусной реальности имени (о количественных параметрах криптоклассного представления имени см. [7, с. 32—38]). Средний показатель *активности имени* по криптоклассу (*ПоКрАИ* 65 сл/соч) может свидетельствовать как о замедлении, так и об ускорении процессов реализации сочетательного потенциала слова (СПС). Однако поскольку индекс периферийности (ИП) имени соответствует нулевой отметке (т.е. имя находится в эпицентре наивных представлений о жидком), очевидно криптоклассные интенции имени полностью реализованы.

Это тот редкий случай, когда удалось обнаружить корпусные свидетельства замещения именем *life* как субъектной, так и объектной (где это возможно) валентности всех десяти глагольных классификаторов криптокласса. Заметим, что корпусные данные о *сочетательной избирательности* имени в данном криптоклассе не отражают его сочетательных пред-

\* Автор благодарит доц. Селезнева Г. Д., доц. Шилихину К. М., проф. Кретова А.А. и выпускников 2009 г. Ситник Н. и Марченко Т. за сотрудничество.

почтений в современном языке. Однако представленные в порядке убывания, такие сведения помогут различить узуальное (>2 сл/соч) VS. окказиональное (<2 сл/соч.) употребление имени: *to flow* 24 (*струиться, течь; хлынуть*), *to pour* 17 (*лить(ся), разливать(ся)*), *to ooze* 7 (*сочиться, вытекать*), *to spill* 7 (*проливать(ся), расплескивать(ся)*), *to stream* 3 (*литься, струиться, течь; источать, лить*), *to leak* 2 (*просачиваться*), *to splash* 2 (*брызгать(ся), плескать(ся)*), *to splatter* 1 (*брызгать(ся), плескать(ся)*), *to sprinkle* 1 (*кропить, накрапывать*), *to flood* 1 (*заливать, затоплять*).

Ср.: *When everything is going fine for us, and God is blessing us, when our life is flowing smoothly, it's easy to have faith, it's easy then to trust... He wants to pour his life into our life so that our life can grow strong. ... Now I'm going to stream Uncle Manfred's life for you, but before I do here are some questions. 'BR. As life flooded back into their daughter, Linda and Junki went limp with relief. ... there are many other ways in which this technology could be used to sprinkle life into Chile's arid zones. ... 'From his canvasses, life spills out'. ... Antonio's family had been mortified by the way his love life was splattered across the papers. ... Oakley's pale as a maiden, the life's leaking from him. ... global superstar Justin Timberlake, has led the low-key actress's private life so that it could be splashed onto the tabloids.*

Рассмотрим имена ближней периферии криптокласса «Liquidus» «Жидкое». При индексе периферийности имен *colour; color* (цвет) и *image* (образ), равном 0,10 (что вполне можно расценивать как ядерный), оба имени не проявляют заметной активности в образовании словосочетаний с глагольными классификаторами данного криптокласса. Интересно, что «цвет, получаемый окрашиванием с помощью pulverизатора» называют *sprinkle* (АВВУ Lingvo 12), однако корпусных свидетельств словосочетания *sprinkle color*\* найдено не было. При этом на фоне слабой активности (ПоКрАИ *color* 71, *image* 47 сл/соч) и окказиональных коллокаций выделяется сочетательная избирательность именами субъектной позиции классификаторов *to flow, to flood*. Ср.: *The bright colors flood my eyes, confusing me. All at once, images and memories flooded his mind in a brilliant whirl, and a revelation blossomed in his soul.*

Вероятно, за этими словами в англоязычном сознании закреплены стойкие ассоциации с **потоком**. Заметим, что и физика цвета, и научная теория отражения исходит «из представлений о **потоковой** природе взаимодействий реальных объектов и механизмов их адаптации» [7, с. 42]. Напомним, что если отражение — универсальное свойство материи, то его суммарным результатом прямой и косвенной деформации при отображении является **образ** (*image*). Как видим, **научная картина мира** ничуть не про-

тиворечит, скорее даже резонирует с наивной (языковой картиной мира). По мере функционирования объекта (например, естественного языка) в роли отражающего его от-**ображ**-ательные способности возрастают по быстродействию и чувствительности, но сужаются по многообразию допустимых праобразов: отсюда диалектическое противоречие между конечностью значений (и знаков в целом) и бесконечностью смыслов [7, с. 43].

Отметим также относительно высокие показатели активности обоих слов в криптоклассах «Rea Parva» и имени *image* в криптоклассах антропоморфного цикла (ПоКрАИ *image* 105, 316, 142 сл/соч). Ожидать реализацию криптоклассных интенций имени *image*, вероятно, следует в криптоклассе «Homo Movens», где наблюдается самое благоприятное соотношение индекса периферийности (0,55) и ПоКрАИ *image* 316 сл/соч. Какими бы не были причины пассивности имен в криптоклассе «Liquidus», можно предположить, что для коммуникации на английском языке образ жидкого образа и цвета является базовым, экспрессивно не окрашенным, связанным с автоматизированным навыком.

Следующая группа имен *figure, information, name* также находится на ближней периферии криптокласса (ИП 0,30). Однако предельно низкий показатель активности *figure* (ПоКрАИ 13 сл/соч) свидетельствует о том, что данная криптоклассная проекция всех фрагментов ментального пространства, покрываемых данной лексемой, не представляет сколько-нибудь заметной значимости для современного англоязычного сознания. По всей вероятности, будучи заимствованным из древнегреческого в значении «форма», слово обменивалось смыслами с другими именами, например, *person, information, digit* и др., вступая с ними в ассоциативные связи по смежности, что предполагало и обмен классификаторами. Ср.: *Perhaps he is referring to some figures that were leaked a couple of weeks ago from the Industrial Development Board for Northern Ireland. Suddenly a darkly handsome figure oozed into the chamber.* Маловероятно образование новых словосочетаний имени с классификаторами данного криптокласса. Заметим, что другие криптоклассные проекции имени представляют большую словотворческую ценность для коммуникации на английском языке.

Имя *information* (ПоКрАИ 298 сл/соч) демонстрирует относительно высокую степень **активности** в криптоклассе. Оно коммуникативно востребовано для выражения смыслов перемещение информации «*information flows*» и нежелательное для говорящего перемещение информации «*to leak information*», «*information leaks*». Ср.: *I certainly hope that the information will be flowing freely to the American people. Information now flows from front-line workers to higher*

management for analysis. They leak information about convoys to members of the Campaign for Nuclear Disarmament... It has been two months since the information leaked out.

На межкриптоклассном портрете (рис.1) видно, что имя *information* представлено на ближней периферии криптоклассов «Liquidus» и «Res Parva» и на дальней периферии криптоклассов антропоморфного цикла.

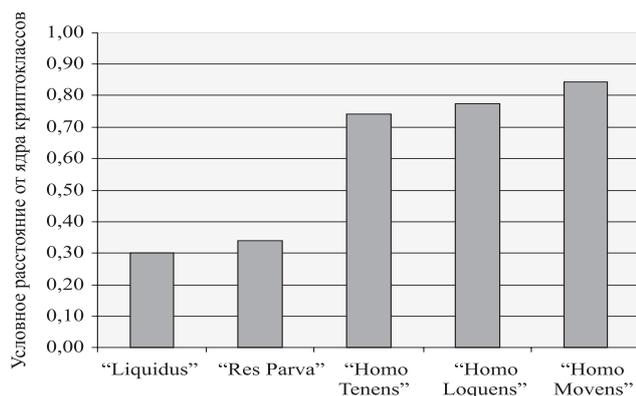


Рис. 1. Сравнение показателей периферийности имени *information* в пяти криптоклассах английского языка

При этом данное имя *максимально активно* в криптоклассе «Res Parva» (рис. 2), и именно с классификаторами этого криптокласса оно, по всей вероятности, будет реализовывать свои криптоклассные интенции. Не менее активным в проявлении криптоклассного статуса жидкого оказалось имя *energy* (ПА 144), которое также имеет четко выраженную **сочетательную избирательность** *energy flows* и *to pour energy*.

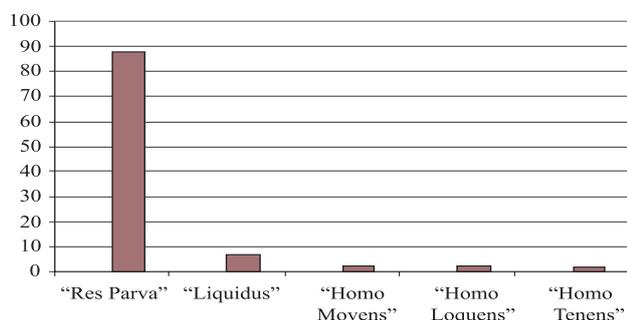


Рис. 2. Относительная активность имени в пяти криптоклассах английского языка

### Эксперимент

Как способ верификации создаваемой модели криптоклассной распределенности имен нами был

предпринят лингвистический эксперимент<sup>2</sup>, основанный на сознательном отношении к решению определенной речевой задачи десяти испытуемых (от 19 до 30 лет, для которых американский английский является родным). Им предлагалось отобрать из предложенных слов 47 глагольных классификаторов криптокласса «Res Parva» — такие, которые могут сочетаться с именем *experience*, и упорядочить их по принципу употребимости (from «most common use» to «unlikely to collocate with»).

Корректировка результатов испытания с данными корпусного анализа сочетаемости имени *experience* (где для каждого глагола посчитана «встречаемость» с именем в текстах) проводилась с двойной целью. Во-первых, проверить, насколько образ мира отдельных испытуемых соответствует картине мира англоязычного сознания. Во-вторых, проверить, насколько реальность естественного дискурса коррелирует с авторским (вероятностным, адаптивным) нормированием сочетаемости.

В результате удалось установить следующее:

1. Списки глаголов всех испытуемых открывают глаголы *to gain*, *to get*, что полностью совпадает с результатами корпусного исследования. См. ниже список глаголов в порядке убывания частоты встречаемости с именем:

*gain* (526), *get* (455), *make* (291), *give* (195), *bring* (175), *take* (101), *put* (35), *capture* (29), *receive* (19), *obtain* (16), *deliver* (15), *gather* (11), *carry* (10), *keep* (9), *place* (9), *pass* (8), *manipulate* (7), *pick* (7), *transfer* (6), *set* (5), *grasp* (4), *hold* (4), *collect* (3), *compress* (3), *cast* (2), *glean* (1), *handle* (1), *seize* (1), *bestow* (0), *brandish* (0), *chuck* (0), *crush* (0), *deposit* (0), *drop* (0), *furnish* (0), *grab* (0), *grip* (0), *hand* (0), *launch* (0), *lay* (0), *precipitate* (0), *press* (0), *shake* (0), *snatch* (0), *squeeze* (0), *throw* (0), *toss* (0).

2. Неожиданно в ряду самых вероятных (по данным девяти анкет) оказались глаголы *to obtain*, *to gather*, *to collect* с относительно низким показателем сочетательной избирательности (СИ) по корпусным данным.

3. Глаголы с относительно высоким показателем СИ (>100) в разном порядке встречаются среди самых или просто вероятных в списках опрошенных.

4. В среднем рейтинг глаголов средней и низкой сочетательной избирательности, например, *to capture*, *to receive*, *to keep* и *to deliver* и др., по результатам эксперимента накладывается на шкалу данных корпусов.

5. У респондентов нет единодушия в том, какие имена следует отнести к категории «unlikely to collocate with». Чаще других здесь можно увидеть глаголы *toss*, *squeeze*, *precipitate*, *throw*, что коррелирует с корпусным анализом сочетаемости имени.

<sup>2</sup> Проведен бакалавром филологии, окончившей СОФ ВГУ в 2009 г. Марченко Татьяной.

6. Выявленные при корректировке несоответствия касаются в первую очередь глаголов с нулевой СИ (т.е. ни одного корпусного свидетельства замещения именем *experience* объектной позиции при глагольных классификаторах криптокласса «Res Parva» не было обнаружено). Однако глаголы *grab (0)*, *launch (0)*, *squeeze (0)*, *drop (0)*, *deposit (0)* в девяти анкетах находятся в середине списка с пометой «not so common, more expressive use».

7. Глаголы с низким показателем СИ: *cast (2)*, *handle (1)*, *seize (1)*, образующие окказиональные словосочетания с именем *experience*, были помечены как «Less common but used», что также соответствует середине списка.

8. Средний показатель СИ глагола *to put (35)* диссонирует с оценкой его сочетательных возможностей у девяти опрошенных: «Uncommon but Maybe».

Таким образом, проведенный нами лингвистический эксперимент показал, что во многом данные корпуса совпадают с интуицией отдельных носителей языка, чьи образы мира представляют собой «репродукции» разной степени точности картины мира англоязычного сознания. Другим важным выводом корреляции данных, полученных разными способами, является перспектива словотворчества. По-видимому, респонденты интуитивно осознают высокую степень вероятности расширения именем *experience* диапазона своей сочетательной избирательности, а именно реализации своих криптоклассных интенций замещать позицию объекта при классификаторах *grab (0)*, *launch (0)*, *squeeze (0)*, *drop (0)*, *deposit (0)*.

И единодушную, достаточно низкую оценку сочетательных возможностей глагола *to put* также следует расценивать как факт изменения именем своих сочетательных предпочтений. Может быть, корреляция вышеописанных методов прольет свет на причины произвольности выбора лексемы?!

#### ЛИТЕРАТУРА

1. Сигал К. Я. Споры вокруг эксперимента в науке о языке и речи / К. Я. Сигал // Психолингвистика в XXI веке : результаты, проблемы, перспективы : тезисы докладов. — М., 2009. — С. 38—47.

2. Стернин И. А. Экспериментальные методы в когнитивной лингвистике / И. А. Стернин // Психолингвистика в XXI веке : результаты, проблемы, перспективы : тезисы докладов. — М., 2009. — С. 415—416.

3. Леонтьев А. А. Слово в речевой деятельности / А. А. Леонтьев. — М., 1965.

4. Мазлумян В. С. Картина мира или образ мира / В. С. Мазлумян // Психолингвистика в XXI веке : результаты, проблемы, перспективы : тезисы докладов. — М., 2009. — С. 82—83.

5. Залевская А. А. Вопросы теории учебного двуязычия и многоязычия / А. А. Залевская // Психолингвистика в XXI веке: результаты, проблемы, перспективы : тезисы докладов. — М., 2009. — С. 297—298.

6. Мельников Г. П. Системология и языковые аспекты кибернетики / Г. П. Мельников. — М., 1978.

7. Борискина О. О. Криптоклассные проекции мира непредметных сущностей : опыт криптоклассного анализа словосочетаемости / О. О. Борискина // Вестник Воронеж. гос. ун-та. Сер. : Лингвистика и межкультурная коммуникация. — 2009. — № 1. — С. 32—38.

Воронежский государственный университет  
Борискина О. О., докторант кафедры теоретической и прикладной лингвистики  
E-mail: olboriskina66@mail.ru  
Тел.: (4732) 20-41-49

Voronezh State University  
Boriskina O. O., Postdoctoral Research Fellow of the  
Theoretical and Applied Linguistics Department  
E-mail: olboriskina66@mail.ru  
Tel.: (4732) 20-41-49