

**Теоретические основания и эксперименты
для анализа соотношения индивидуальных словарей с лексической
системой языка**

1. Проблема соотношения индивидуальных словарей с лексической системой языка.

Анализу индивидуального языкового знания до сих пор не везёт. Рассуждения о языке «в целом» редко опосредуются анализом того, **а как, какими частями, в какой степени** это целое представлено в индивидуальных мозгах (не говоря уже об анализе целостного механизма действия индивидуального словаря). Даже определение такой относительно простой характеристики, как объём индивидуального словарного запаса индивидуума, до сих пор ориентируется на умозрительно-фантастические оценки, находящиеся в очень широком диапазоне - от нескольких сотен слов до нескольких сотен тысяч (Денисов, 1969). В то же время прямые эксперименты по тестированию русских индивидуумов со средним и незаконченным высшим образованием (Савчук, 1997) и даже неграмотных крестьян (Тимофеев, 1971) дают не меньше нескольких десятков тысяч слов, известных рядовому носителю языка. В некоторых экспериментах с испытуемыми с высшим образованием средние и максимальные величины знания русского литературного словаря поднимаются ещё выше – до 100 тысяч и более слов (Поликарпов, Поликарпова, 1987). Однако подобных проведённых в ограниченном масштабе исследований совершенно недостаточны. Они пока не дают определённого ответа ни на вопрос о пределах индивидуального знания словаря (т.к. используемые тестовые списки явно не являются максимальными и, значит, не позволяющими оценить объём индивидуального словаря сверху), ни на вопрос о соотношении в нём активной и пассивной его частей, ни на то, как

формируется и развивается лексическая система в целом, во всей совокупности индивидуальных голов. Для продвижения в этом направлении необходимо, во-первых, подвергнуть теоретическому анализу вопрос о базовом механизме взаимодействия общего словаря данного языкового обществ и индивидуальных словарей его членов. Во-вторых, необходимо расширить на этой основе эмпирическую базу исследования этих вопросов.

Излагаемая ниже концепция соотношения и взаимодействия индивидуальных словарей с лексической системой языка базируется на работах членов лаборатории общей и компьютерной лексикологии и лексикографии (Поликарпов, 1987; 1998; 2011; 2012 - в печати; Поликарпов, Поликарпова, 1987; Поликарпов, Крюкова, 1989; Поликарпов, Курлов, 1994; Савчук, 1997).

В своём моделировании словаря (и языка в целом) как социального явления мы, вслед за указанными работами, исходим из того, что исходной реальностью существования языка являются индивидуальные словари. Каждый индивидуальный словарь уникален и является результатом общения данного коммуниканта в различных социальных группах (семейных, профессиональных, спортивных и т.п.). В ходе коммуникации происходит постоянное **сопоставление** индивидуальных словарей (с выявлением совпадений и расхождений в них) и **взаимоуподобление** (усиление повторяющихся черт и ослабление различных). Причём, сопоставление происходит на фоне членения каждого из индивидуальных словарей на две принципиально различные части - **АКТИВНУЮ** и **ПАССИВНУЮ**. Активная часть индивидуального словаря – это то, что используется носителем из всего этого словаря в своей **продуктивной фазе общения** – при говорении и писании в адрес партнёра по общению. Пассивная часть индивидуального словаря – это то, что используется носителем из всего этого словаря только в своей рецептивной фазе общения – при чтении и слушании текстов, обращённых к нему. Причём, предположительно, **активная часть индивидуального словаря в норме**

всегда меньше по объёму, чем пассивная. Это объясняется тем, что пассивная часть формируется включённостью члена общества в большее число разных актов общения с разными людьми разных индивидуальных и групповых интересов (и поэтому восприятием и отображением в каждом пассивном индивидуальном словаре большего числа разных употреблённых другими слов), чем это происходит при формировании активной части его индивидуального словаря. Активная часть словаря формируется более узкой, менее разнообразной продуктивной активностью каждого индивидуума в коммуникации.

При принципиальной уникальности состава каждой из активных частей индивидуальных словарей в данном коллективе обнаруживается и существенное их **ПЕРЕСЕЧЕНИЕ**, значимое **тождество** по набору входящих в них слов и их значений. Это пересечение – не что иное, как тот **социально значимый лексический минимум каждого полноценного члена общества, создаваемый и поддерживаемый их взаимным общением-обучением.** Этот тот минимум, который покрывает, видимо, наиболее важные, активные общезначимые, ядерные смыслы жизни данного общества. Это тот минимум, на основе которого можно **объяснить (описать) любой другой смысл, который выходит за пределы ядерного их набора.** Т.е. те лексические единицы, которые входят в его состав, образуют тот **фундамент, на котором базируется общение.** Это пересечение – наиболее важная часть общего словаря языка. Это – **ЯДРО, базовый, основной слой общего словаря языка,** единицами которого пользуются **обязательно все полноценные члены данного общества.** Это и тот обязательный минимум, которым должен овладеть и любой некоренной носитель данного языка, чтобы стать полноценным его носителем. Это и тот круг лексики, от которого «сверху» должны исходить методисты при разработке менее объёмных лексических минимумов различного рода (чего в действительности практически не наблюдается).

Кроме того, это и тот круг лексики, который, как показывают теоретико-лексикографические исследования, является основным содержанием **словаря краткого типа**, который занимает **центральное положение в системе толковых словарей любой национальной языковой культуры**. Дело в том, что наиболее авторитетные толковые словари, с теми или иными оговорками (см., например, [Поликарпов, 1987; Поликарпов, Крюкова, 1988; Поликарпов, Курлов, 1994; Малов, 1986; Обухова, 1986]) являются тем или иным специфическим системным типизированным отображением реально существующей языковой действительности – отображением того или иного слоя (или набора слоёв) лексической системы. Исследования показывают, что, при имеющей место известной субъективности в работе лексикографов, отбор слов, значений и их характеристик (в том числе и стилистических) в словари разного типа происходит совсем не случайным образом, а на основе определенных задаваемых, "навязываемых" объектом принципов, что приводит к системной, закономерной соотнесенности состава толкового словаря с общим составом лексической системы языка. Эти принципы задаются объективным устройством языковой действительности, и лексикографы эти принципы не придумывают, а лишь в той или иной мере учитывают, отображают. Таким образом, можно говорить о том, что словарь - не хаотическое нагромождение фактов, что в словарях, вероятно, определенным образом отражаются реально существующие в языке пласты и сферы словарного состава языка, некоторые характерные особенности его системного устройства.

Идея учёта того, как различные по своей ценности слои лексики отражаются в составе разных толковых словарей, была впервые осознана и реализована в оригинальной **классификации типов толковых словарей**, предложенной в 1952 году С.И.Ожеговым. Основным, базовым в его классификации является **толковый словарь краткого типа** [Ожегов, 1952]. Он им был определён как *«краткий, популярного типа,*

стремящийся к активной нормализации современной литературной речи" [1952, с. 91-92]. Кроме этого, самого базового типа толковых словарей им выделяется ещё и тип **среднего толкового словаря**, «с детальной разработкой исторически оправданного стилистического многообразия современного литературного языка», а также тип **большого толкового словаря** «представляющий современный литературный язык в широкой исторической перспективе» [1952, с. 91-92]).

Действительно, если выделенный С.И.Ожеговым тип краткого словаря отражает ядро, т.е., видимо, самое важное в лексике всего языка, то словари такого типа должны являться самыми нужными, самыми популярными и должны в наиболее строгой степени ориентироваться на сохранение, нормализацию лексического материала, отражённого в нём. Какие реальные словари (если они есть) представляют данный тип национального словаря-минимума в русской лексикографической традиции? В русской лексикографической традиции толковый словарь краткого типа представлен различными изданиями «Словаря русского языка» С.И.Ожегова [Ожегов, 1949 -...].

Ниже эта серия словарей будет дополнительно охарактеризована. Пока же зададимся вопросом, насколько универсален, насколько обязательно представлен этот тип конкретными национальными его вариантами в других национальных языковых культурах? Может быть, выработка такого типа и всей иерархии типов толковых словарей - это только специфика русской национальной культуры? На самом деле, как оказывается, краткий тип словарей, как и остальные типы (среднего и большого) словарей представлен в любой развитой лексикографической традиции. Что и неудивительно, т.к. принципиальная структура словарного сознания общества на определённом (одном и том же для сопоставляемых культур) этапе их развития едина, т.к. формируется одним и тем же, универсальным для них набором факторов.

Далее, надо осознать, что кроме наиболее ясно выделяющегося «ядра» лексики, образуемого **ПЕРЕСЕЧЕНИЕМ** индивидуальных активных словарей, достаточно объективным (хотя и менее точно определяемым) онтологическим статусом в обществе обладает и **ОБЪЕДИНЕНИЕ** этих активных словарей. Это «объединение» есть не что иное, как **совокупный актуальный лексический инвентарь данного общества в данное время**. Т.е. это полная сумма того лексического материала, который используется кем-либо из ныне живущих людей в данном языковом коллективе. Совершенно очевидно, что подобный объект менее обозрим и менее ощутим, чем пересечение индивидуальных активных словарей. Однако подобный объект тоже является предметом пристального внимания в каждой развитой национальной лексикографической традиции, т.к., совершенно очевидно, что в развитых обществах нового времени присутствует и задача фиксации и нормализации всей социальной лексической практики, **всего актуального лексического материала языка**, а не только его «ядра».

Дополнение к «ядру» лексики, выделенному в результате практического приложения операции пересечения к общей совокупности индивидуальных активных словарей в этом случае представлено в наибольшей степени периферийными, специальными областями общения (специально-научными, профессионально-техническими, социально-групповыми и т.п.). Справочно-нормирующая информация, фиксируемая в словарях о словах и значениях этих областей, хоть и менее социально значима, чем информация о лексическом «ядре», но тоже необходима обществу на достаточно продвинутых этапах его развития, т.к. специальные области – это точки роста культуры. А кроме того, центр и периферия лексической системы находятся в сложном взаимодействии, из которого исключить одну из этих сторон невозможно без потери перспектив понимания закономерностей существования целого механизма этой системы.

Тип словарей, отображающих "активный и пассивный запас лексики современного языка", представляющий «стилистическое многообразие современного языка» был определён С.И.Ожеговым как **тип «среднего» словаря** [Ожегов, 1952, с. 92, 100]. В русской лексикографической традиции к словарям "среднего" типа относится, кроме МАС'а двух изданий [Евгеньева, 1957-1961; 1981-1984], еще "Толковый словарь русского языка" под ред. Д.Н. Ушакова [Ушаков, 1938-1940].

Завершая принципиальную структуризацию лексической системы языка, необходимо осознать, что каждый индивидуальный словарь обладает ещё и **пассивной частью**, в которую входят помимо известных некоторому индивидууму, но неупотребляемых им (хотя и употребляемых актуально кем-то другими в данном обществе в данное время) современных слов и значений (в основном, специально-периферийного плана, в основном, профессионально-терминологические единицы), ещё и различного рода **устаревшие** единицы, известные ему, в основном, из письменных текстов, относящихся к «современному литературному языку» в широком смысле – «от Пушкина до наших дней». Это **дополнение** к общей сумме, к объединению разных по активности и распространённости актуальных лексических единиц (отображаемых, как было сказано, в словарях среднего типа), состоящее только из устаревших, реально нигде не представленных, кроме текстов прежних эпох, продолжающих сохранять культурную значимость единиц и приводит к **ПОЛНОМУ ОХВАТУ ЛЕКСИЧЕСКОЙ СИСТЕМЫ ЯЗЫКА**. Устаревшее «дополнение» представляет тот слой лексической системы, который является для неё наиболее периферийным, но всё же реально существующим, т.к. хотя бы некоторая часть общества хотя бы и пассивно, но знает единицы этого слоя, читает тексты, содержащие их.

В наиболее полном виде (в совокупности всех трёх указанных слоёв) лексическая система отображается составом толкового словаря, называемого в лексикографической традиции **«большим»** [Ожегов, 1952].

К "большим" словарям в современной русской лексикографии относится, прежде всего, "Словарь современного русского литературного языка" в 17-ти томах [БАС, 1948-1965]. В последние годы были изданы сопоставимые с ним по объёму словника и набора значений такие словари, как «Большой толковый словарь русского языка» [Кузнецов, 1997] и «Новый словарь русского языка» [Ефремова, 2001].

В итоге, как видно, состав словаря (набор слов и их значений) краткого типа, отражая наиболее ядерную, наиболее активную, самую нужную в любом акте общения лексику, должен обладать полной вложимостью в словарь среднего типа (содержащий, в принципе всю нормативно отобранную современную лексику), а состав среднего словаря должен обладать полной вложимостью в большой словарь (содержащий, в принципе, всю нормативную лексику данного языка за весь длящийся период «современного языка»), причем, и по словнику, и по набору значений каждого из повторяющихся слов в этих словарях. Это представление о соотношении словников и набора значений толковых словарей разных типов, в принципе, подтверждается, например, при соотнесении количественных характеристик полисемических структур русских словарей всех трёх типов и английских толковых словарей краткого и среднего типов (Поликарпов, 1987), а также при поэлементном соотнесении словника и значений двух русских толковых словарей – словаря краткого типа («Словаря русского языка» С.И.Ожегова, 9 изд., 1972 г.) и словаря среднего типа («Словаря русского языка в 4 тт.» под ред. А.П.Евгеньевой, 1981-1984) [Поликарпов, Крюкова, 1989]). Это представление подтверждается и прямыми тестами типов словарного знания коренных носителей русского языка, которым посвящена настоящая статья.

Исходя из подобных представлений, можно понять, что наибольшей важностью для любой национальной языковой практики обладает лексический материал, отображаемый в словаре краткого типа. В кратком

словаре должны концентрироваться ядерные для языка слова и их значения, т.е. те, которые **известны всем и используются всеми полноценными носителями языка**. Это тот материал, который обладает наибольшей ценностью для взаимопонимания в любом языковом коллективе. Этот пласт лексики постоянно вероятностно формируется в ходе взаимного «плюс» и «минус» обучения носителей языка в практических актах коммуникации. Каждый говорящий в общении ощущает, что из его слов сразу понятно, а какие слова надо заменить другими, чтобы на их основе разъяснить другие, сразу не понятные слова (или их значения). Кроме того, он может ощущать и в речи партнёров по общению усиление активности употребления каких-то известных ему слов или их значений, что может заставить и его ввести эти лексические единицы в свой активный лексикон. Это и формирует в языковом коллективе тот вероятностно ощущаемый каждым стандарт, который мы назвали **«пересечение индивидуальных активных словарей»**.

«Представители» этого ядерного, наиболее важного слоя, по нашим предположениям, могут быть надёжно представлены в материалах тестирования индивидуального словарного знания, если отталкиваться от тестовых списков, базирующихся на больших, представительных для всего языка в целом словарях русского языка. Таковым, наиболее полным по своему составу на сегодняшний день, является «Сводный словарь современной русской лексики», содержащий ок. 165 тыс. разных слов, собранных из разнообразных лексикографических источников – толковых, терминологических, энциклопедических и иных словарей. Отобранный нами случайным образом из него словник в 1651 слово может служить надёжным представителем всего словаря. Этот тестовый словник подобен тем выборкам населения данной страны, которые производятся при социологических опросах. Он так же, как и представительная социологическая выборка, вероятностно соответствует полному объекту. По нему можно получить достаточно точную и объективную информацию

об основных его характеристиках (полисемических, омонимических, стилистических, семантических и т.п.)

. 2. Методика и материал исследования.

Материалом для дящегося несколько лет исследования послужили результаты прямого эксперимента по тестированию словарного знания носителей русского языка одной возрастной и образовательной группы (студенты 3-5 курсов 20-25 лет филологического ф-та МГУ). К настоящему времени проведён пилотный анализ начальной части этих данных - для 20 испытуемых [Поликарпов, 2011; 2012 – в печати]. В настоящем исследовании представлены результаты анализа этой группы.

Список слов для тестирования был составлен на основе равномерно-случайной выборки (каждое очередное 100-ое слово) из «Сводного словаря современной русской лексики» под ред. Р.П.Рогожниковой (1990) общим объёмом в 165 тыс. слов. Выбор этого словаря в качестве источника данных объясняется тем, что это наиболее обширное собрание лексики, полученное путём объединения словников (и основной грамматической и омонимической информации, а также информации о вариативности облика слов) из основных толковых, терминологических и энциклопедических словарей современного русского языка.

Ниже приводится список словарей, входящих в состав "Сводного словаря современной русской лексики":

- У - Толковый словарь русского языка / Под ред. Д.Н. Ушакова. М., 1935 - 1940
- У(д) - Дополнения к первым трем томам // Толковый словарь русского языка / Под ред. Д.Н. Ушакова. М., 1940. Т. 4
- М, МАС - Словарь русского языка: В 4 т./ АН СССР. Ин-т рус. яз. М., 1957 - 1961. Т. 1 - 4
- Б, БАС - Словарь современного русского литературного языка/ АН СССР. Ин-т рус. яз. М., Л., 1948 - 1965. Т. 1 - 17
- И - Словарь иностранных слов / Под ред. И.В. Лехина, С.М. Локшиной, Ф.Н. Петрова (гл. ред.), Л.С. Шаумяна. 6-е изд., перераб. и доп. М., 1964. 784 с.
- НС - Новые слова и значения: Словарь-справочник по материалам прессы и литературы 60-х годов / Сост.: Н.З. Бутарова, Н.З. Котелова (руководитель группы), Е.А. Левашов и др. / Под ред. Н.З. Котеловой, Ю.С. Сорокина; АН СССР. Ин-т рус. яз. М., 1971, 544 с.
- Ож, О, СО - Ожегов С.И. Словарь русского языка / Под ред. Н.Ю. Шведовой. 10-е изд., стер. М., 1973, 848 с.
- Орф - Орфографический словарь русского языка / Под. ред. С.Г. Бархударова, И.Ф. Протченко, Л.И. Скворцова; АН СССР. Ин-т рус. яз. 13-е изд., испр. и доп. М., 1974, 480 с.
- П - Политехнический словарь / Гл. ред. И.И. Артоболевский; редкол.: Д.М. Алексеев, Д.М. Беркович, В.Г. Воскобойников и др. М., 1976. 680 с..

БСЭ - Большая Советская Энциклопедия / Гл. ред. А.М. Прохоров. 3-е изд. М., 1969 - 1978. Т. 1 - 30 И - 7 - Словарь иностранных слов / Под ред. А.Г. Спиркина, И.А. Акчурина, Р.С. Карпинской. 7-е изд., перераб. М., 1979. 264 с.

СЭС - Советский энциклопедический словарь / Научно-редакц. совет: А.М. Прохоров (председатель), М.С. Гиляров, Е.М. Жуков и др. М. 1979. 1600 с.

М-2, МАС-2 - Словарь русского языка: в 4 тт. / Гл. ред. А.П. Евгеньева; АН СССР. Ин-т рус. яз. 2-е изд., испр., доп. М., 1981 - 1984. Т. 1 - 4

ОЭ - Орфоэпический словарь русского языка: Произношение, ударение, грамматические формы / С.Н. Борунова, В.Л. Воронцова, Н.А. Еськова; Под ред. Р.И. Аванесова; АН СССР. Ин-т рус. яз. М., 1983. 704 с.

НС-2 - Новые слова и значения: Словарь-справочник по материалам прессы и литературы 70-х годов / Е.А. Левашов, Т.Н. Поповцева, В.П. Фелицына и др.; Под ред. Н.З. Котеловой; АН СССР. Ин-т рус. яз. М., 1984. 808 с.

Полученный список содержит 1651 слово, каждое из которых каждый из участников эксперимента классифицировал по предложенным в анкете разделам, исходя из своих представлений об УПОТРЕБЛЕНИИ, ПАССИВНОМ ЗНАНИИ ИЛИ ПОЛНОМ НЕЗНАНИИ каждого тестового слова в собственной речевой практике.

Полученный в ходе эксперимента исходный материал тестирования каждого испытуемого подвергся контролю (через анализ кратких толкований, дававшихся испытуемыми). Полученные таким образом данные вместе с общеязыковыми данными по каждому из 1651 слова (величина полисемии, стилистические пометы, величина омонимии, представленность в каждом из 4 указанных словарей и т.п.), были внесены в базу данных (на основе электронных таблиц Excel – 2007), которая представляет собой матрицу 1651x95.

БД содержит следующие поля (в виде столбцов):

- слово,
- его часть речи,
- сводная полисемия всех омонимов данного слова,
- степень омонимичности слова,
- стилистические пометы значений слова,
- денотативная область значений слова,
- встречаемость слова в СО,
- встречаемость слова в МАСе,
- встречаемость слова в Ушаковском словаре,
- встречаемость слова в БАСе,
- полная «формула» встречаемости слова в указанных выше словарях,
- набор значений данного слова по БАСу.

Для каждого из 20 испытуемых выделены 5 полей, характеризующих их знание очередного слова:

- «УПОТРЕБЛЯЮ»,
- «ЗНАЮ, НО НЕ УПОТРЕБЛЯЮ»,
- «НЕ ЗНАЮ»,
- краткое толкование хотя бы одного значения слова,
- оценка экспериментатором правильности толкования,

3. Анализ некоторых полученных результатов

Анализ базы данных заключался в анализе характеристик самого списка, а также в составлении частотного списка слов по количеству активных и негативных словарей испытуемых, в которых они представлены, а также в попарном и последовательном полном соотнесении активных и негативных словарей разных испытуемых, выявлении их пересечений, объединений и т.п., выявлении особенных системных языковых характеристик пересечения активных словарей и пересечения негативных словарей испытуемых.

3.1. Общие характеристики тестового списка.

Таблица 1. Распределение слов тестового списка по частям речи

Части речи	количество слов
сущ	826
глагол	461
прил	248
нар	104
межд	6
част	2
числ	2
вводное слово	1
предл	1
Общий итог	1651

Данная таблица даёт самое общее представление о соотношении слов разных частей речи в большом словаре. Ок. половины всего списка составляют существительные. Ввиду выборочности наших данных некоторые малочисленные служебные части речи не получили отображения в нашем списке.

Таблица 2. Распределение слов тестового списка по величине многозначности

МНОГОЗНАЧНОСТЬ	1	2	3..4	5..8	9...16	Общий итог
К-во слов д. вел. мн-ти	778	492	29 0	7 8	9	1651

Как видно, в распределении слов по числу значений большую половину составляют малозначные (1-2-значные) слова.

3.2. Соотношение величин активных, пассивных и полных индивидуальных словарей

Для каждого из 20 испытуемых были исходно выявлены:

(1) АКТИВНЫЙ ИНДИВИДУАЛЬНЫЙ СЛОВАРЬ,
 (2) ПАССИВНЫЙ ИНДИВИДУАЛЬНЫЙ СЛОВАРЬ,
 (3) ПОЛНЫЙ ИНДИВИДУАЛЬНЫЙ СЛОВАРЬ (АКТИВНЫЙ + ПАССИВНЫЙ),

(4) НЕГАТИВНЫЙ ИНДИВИДУАЛЬНЫЙ СЛОВАРЬ.

Соотношение их объёмов представлено ниже в **Таблице 3**.

Таблица 3. Соотношение испытуемых по величинам их активных, пассивных, полных и негативных индивидуальных словарей

Испытуемые	Слова ри испыт уемых	Активный	Пассивный	Полный	Негативный
		Алк.	842	200	1042
Бур.	595	652	1247	404	
Вол.	769	531	1300	351	
Дом.	656	489	1145	506	
Дыб.	1086	293	1379	272	
Лап.	840	480	1320	331	
Мор.	685	522	1207	444	
Пав.	692	409	1101	550	
Ста.	753	471	1224	427	
Стр.	595	657	1252	399	
Фис.	694	210	904	747	
Шал.	687	604	1291	360	
Кол.	664	241	905	746	
Мих.	850	444	1294	357	
Ара.	705	502	1207	444	
Ток.	584	661	1245	406	
Дос.	709	590	1299	352	
При.	684	416	1100	551	
Сел.	772	451	1223	428	
Мар.	830	485	1315	336	

Среднее	734	465	1199	452
---------	-----	-----	------	-----

Из приведённой таблицы видно, что **величина активного словаря** испытуемых колеблется в пределах от **584 до 1086 слов**.

Величина пассивного словаря испытуемых колеблется от **200 до 661 слов**.

Величина полного словаря испытуемых колеблется от **904 до 1379 слов**.

Из полученных данных следует, что величина активного словаря студентов филологов старших курсов МГУ (молодых людей с незаконченным высшим образованием) в расчёте на весь язык чаще всего не опускается ниже 60.000 слов (если пересчитывать объем тестового активного списка на полный объем «Сводного словаря современного русского языка», откуда случайным образом отбирался тестовый список). Величина же полного словаря (активный плюс пассивный) в подавляющем большинстве случаев превышает 100.000 слов (по той же самой пропорции).

Отметим, что как оценки активных, так и пассивных словарей – это оценки «снизу». Эти оценки активных, пассивных и полных словарей являются вполне реалистическими, они дают представление о том, **НЕ НИЖЕ КАКОГО ОБЪЁМА** индивидуальные (активные и, в особенности, пассивные) словари у наших испытуемых. Реалистичность этих данных, во-первых, определяется тем, что каждая из квалификаций знания того или иного слова испытуемыми при обработке результатов тестирования верифицировалась и подтверждалась экспериментаторами. Во-вторых, это определяется тем, что в качестве источника тестового материала выступало самое большое из опубликованных собраний лексики – «Сводный словарь современной русской лексики» (Рогожникова, 1990). В-третьих, это определяется случайным способом отбора тестового списка лексики из данного словаря – отбиралось каждое очередное 100 слово словника этого словаря. При условии, что алфавитный порядок, в котором стоят слова в словаре, является, в основном, случайным для существенных свойств его слов, этот принцип отбора позволяет получить **ПРЕДСТАВИТЕЛЬНЫЙ** список для этого словаря (и, видимо, для языка в целом).

Вместе с тем, то, что пассивный словарь личности, по данным этих тестов, оказывается в подавляющем числе случаев **меньше**, чем соответствующий активный словарь, говорит о том, что тестовый список отображает состав словника словаря, явно не полного по набору слов. Как для активного, так и, в большей степени для пассивного словаря полученные нами оценки являются пока оценками «снизу», оценками того, не ниже какого объёма эти словари. Более полные оценки как активного, так и пассивного словарей носителей русского языка смогут быть получены в будущем на основе формирования и использования в качестве генеральной совокупности более полных сводных русских словников.

3.3. Попарные и последовательные пересечения и объединения единиц активных и пассивных идиолексиконов испытуемых.

Определение величин всех возможных попарных и последовательных пересечений и объединений слов активных и пассивных словарей испытуемых позволяет получить исходные данные для анализа общей принципиальной картины возможного соотношения индивидуальных словарей в языковом коллективе.

Ниже приводится **Таблица 4**, содержащая фрагмент общего пространства соотношений активных словарей испытуемых и их пересечений.

Таблица 4. Величины активных словарей испытуемых и их попарные пересечения (фрагмент общей таблицы)

Словарь1	Пересечение 1-2	Словарь2	Словарь1	Пересечение 1-2	Словарь2
	Алк/Бур			Бур/Вол	
Алк	Пересечение	Бур	Бур	Пересечение	Вол
842	507	595	595	480	769
Алк / Вол	Бур / Дом				
Алк	Пересечение	Вол	Бур	Пересечение	Дом
842	638	769	595	430	656
Алк / Дом	Бур / Дыб				
Алк	Пересечение	Дом	Бур	Пересечение	Дыб
842	547	656	595	535	1086
Алк / Дыб	Бур / Лап				
Алк	Пересечение	Дыб	Бур	Пересечение	Лап
842	755	1086	595	505	840
Алк / Лап	Бур / Мор				
Алк	Пересечение	Лап	Бур	Пересечение	Мор
842	665	840	595	433	685
Алк / Мор	Бур / Пав				
Алк	Пересечение	Мор	Бур	Пересечение	Пав
842	544	685	595	455	692

Можно заметить, что попарное пересечение активных словарей редко содержит в себе менее 50% слов от каждого из сопоставляемых индивидуальных активных словарей. Это свидетельствует о том, что на уровне индивидуальных активных словарей полноценные носители современного русского языка (одной и той же социальной группы) весьма близки друг к другу. Наличие высокой степени сходства между активными индивидуальными словарями определяет возможность высокой степени взаимопонимания в общении для носителей русского языка.

На основе наших данных можно проследить динамику сужения площади пересечения индивидуальных активных знаний словаря при

увеличении числа сопоставляемых индивидуальных активных словарей испытуемых – от 2 до 20 человек (см. **Таблицу №5** и **Рис. № 1**).

Таблица №5. Динамика сужения площади пересечения индивидуальных активных словарей при последовательном увеличении числа испытуемых в группе

Объёмы групп испытуемых	Число слов в пересечении индивидуальных словарей каждой группы
1	842
1...2	507
1...3	440
1...4	363
1...5	356
1...6	333
1...7	276
1...8	256
1...9	242
1...10	226
1...11	214
1...12	214
1...13	212
1...14	212
1...15	212
1...16	212
1...17	210
1...18	210
1...19	210
1...20	210

Из этого рисунка, в частности, следует, что последовательное уменьшение величины общего активного словаря при росте числа сопоставляемых словарей постоянно замедляется, доходя до стадии стабилизации его величины на уровне 210 слов, т.е. на уровне реального индивидуального словаря величиной **ок. 20 тысяч слов**.

Разумно предполагать, что эта величина близка к реальной предельной, т.е. к той, которая может быть получена на той же категории носителей русского языка, но с её дальнейшим количественным расширением. Более того, эта величина близка к предельной и в том случае, если бы мы

использовали более объёмный тестовый список. Дело в том, что расширение словаря сводного словаря происходит, в основном, за счёт добавления в список лексических единиц всё более периферийных областей словаря – специальных, регионально маркированных, устарелых, т.е. тех, которые практически не представлены в активных индивидуальных словарях и поэтому принципиально не представлены и в пересечениях этих активных словарей.

Аналогично моделируется динамика сужения площади пересечения индивидуальных пассивных словарей при росте числа сопоставляемых словарей (см. **Таблицу №6** и **Рис. №2**).

Таблица №6. Динамика сужения площади пересечения индивидуальных пассивных словарей при последовательном увеличении числа испытуемых

Группы испытуемых	Число слов в пересечении их пассивных словарей
1	200
1...2	135
1...3	88
1...4	49
1...5	18
1...6	13
1...7	7
1...8	4
1...9	3
1...10	3
1...11	0
1...12	0
1...13	0
1...14	0
1...15	0
1...16	0
1...17	0
1...18	0
1...19	0
1...20	0

Рис. №2. Динамика уменьшения числа слов в пересечении пассивных индивидуальных словарей при увеличении числа испытуемых

Как видно, здесь пределом пересечения этого типа словарей является нулевой «общий пассивный словарь». Т.е. разнообразие индивидуальных пассивных словарей существенно выше, чем активных, что и даёт даже при небольшом числе испытуемых быстрое сокращение объёма их пересечения при увеличении объёма группы испытуемых и неизбежный приход этого пересечения к нулю.

Кроме того, необходимо отметить, что индивидуальные активные словари (в отличие индивидуальных пассивных) находятся в интенсивном взаимодействии, что и ведёт к их постоянному взаимоуподоблению (препятствующему их «разбеганию» под влиянием разнообразия условий и целей их использования своего языка), к поддержанию определённой степени общности активной части индивидуальных словарей в обществе. Их пересечение никогда не может быть нулевым, т.к. активные индивидуальные словари – основа взаимопонимания членов общества.

4.4. Выявление зависимости специфических общеязыковых характеристик слов тестового списка от числа испытуемых, употребляющих их в разной степени активно и не употребляющих их вовсе

В приводимых ниже **Таблице 7** и **Рис. 3** представлены данные по общему распределению числа слов тестового списка в зависимости от их представленности в большем или меньшем числе активных индивидуальных словарей. Слов, которые знают все 20 испытуемых – 210. Однако тех слов, которые не знает никто, ещё больше – 224.

Таблица № 7. Общее распределение числа слов в зависимости от степени их представленности в активных идиолексиконах

Активность слов (число испытуемых, употр. слова активно)	Ск. таких слов
0	224
1	152
2	92
3	72
4	69
5	60
6	60
7	56
8	55

9	52
10	46
11	44
12	38
13	38
14	42
15	47
16	50
17	54
18	58
19	132
20	210

Эти данные обнаруживают явный факт двувершинности данного распределения. Левая вершина определяется, видимо, тем фактором, что определённое число слов в тестовом списке относится к устойчиво-устарелому слою в языке. Чем большее число активных индивидуальных словарей пересекается, тем меньше в этом пересечении оказывается таких слов. Последовательное сокращение числа слов в таком пересечении на правой части распределения останавливается, видимо, действием того фактора, что какое-то значительное число слов (в тестовом списке их 210, что в пересчёте на полный словарь даёт ок. 20 тыс. разных слов) в результате постоянного взаимного языкового обучения членов общества оказывается устойчиво употребляемыми всеми носителями языка данной био-социальной группы.

Кроме того, анализ показывает, что **чем меньше наших испытуемых (из общего количества 20 испытуемых) содержат те или иные слова в их активном словаре, тем эти слова менее полисемичными**. Особенно немногочисленны слова из негативного словаря (списка слов, которые никто из испытуемых не знает).

Что касается **отличий слов разных частей речи**, то для **существительных** (наиболее предметной части речи) характерной является наибольшая концентрация в общем негативном словаре испытуемых. А для наречий (наиболее признаковой части речи среди знаменательных слов) – концентрация в области самых активных слов.

На основе тезаурусной схемы, предложенной в «**Тезаурусе синонимов современной русской речи**» под ред. Л.Г.Бабенко (2007), в нашей базе охарактеризованы семантические категории значений слов тестового списка, в том числе расклассифицированы значения 210 слов, которые представлены

в пересечении активных словарей испытуемых, а также значения 224 слов, которые не употребляет никто из наших испытуемых.

Ниже приводятся **Таблицы № 8 и 9**, содержащие в себе данные по частотности (степени представленности) различных семантических категорий значений для самых активных 210 слов (употребляемых всеми) и самых неактивных 224 слов (никем не употребляемых). Это сопоставление принесло ещё более интересные результаты.

Таблица 8. Распределение по различным семантическим категориям процентов значений 210 слов тестового списка, употребляемых всеми 20 испытуемыми

Семантические категории	Проценты значений слов каждой сем. категории среди всех значений группы слов, входящих в пересечение активных идиолексиконов
10. Социальная деятельность.	21
15. Восприятие окружающего мира.	15
16. Универсальные представления, смыслы и отношения.	13
4. Эмоции	13
2. ЖИВАЯ ПРИРОДА	11
12. Быт	5
11. Социальная сфера жизни человека.	5
1. НЕЖИВАЯ ПРИРОДА	4
7. Интеллект	3
6. Речь	3
5. Оценка	2,6
9. КОНКРЕТНАЯ ФИЗИЧЕСКАЯ ДЕЯТЕЛЬНОСТЬ	2

3. Человек как живое существо	2
13. Населенный пункт	0,4
Общий итог	100

Таблица 9. Распределение по различным семантическим категориям процентов значений 224 слов тестового списка, не употребляемых никем из 20 испытуемых

Семантические категории	Проценты значений слов каждой сем. категории среди всех значений группы слов, входящих в пересечение негативных идиолексиконов
10. Социальная деятельность.	27
2. ЖИВАЯ ПРИРОДА	14
1. НЕЖИВАЯ ПРИРОДА	11
12. Быт	9
11. Социальная сфера жизни человека.	8
16. Универсальные представления, смыслы и отношения	8
3. Человек как живое существо	6
6. Речь	6
4. Эмоции	4
15. Восприятие окружающего мира	3
14. Нации	2
7. Интеллект	2
Общий итог	100

Как видно из этих таблиц, крайняя периферия современного словаря, а именно, слова, которые не знает никто из испытуемых, характеризуется в большей степени предметными номинациями (**ЖИВАЯ ПРИРОДА, НЕЖИВАЯ ПРИРОДА, БЫТ**), хотя на 1-ом месте по активности стоит категория **СОЦИАЛЬНАЯ ДЕЯТЕЛЬНОСТЬ**.

В то же время активное ядро словаря русского языка концентрируется (в порядке падения степени представленности) вокруг совершенно очевидно «непредметных» категорий:

- **15. ВОСПРИЯТИЕ ОКРУЖАЮЩЕГО МИРА;**
- **16. УНИВЕРСАЛЬНЫЕ ПРЕДСТАВЛЕНИЯ, СМЫСЛЫ И ОТНОШЕНИЯ;**
- **4. ЭМОЦИИ.**

Полученное в нашем исследовании пересечение всех активных словарей испытуемых (**210 слов**), по всей видимости, относится к ядру лексической системы русского языка. Эти «ядерные» свойства этого списка особенно хорошо видны на фоне свойств того списка из **224 слов**, которые не употребляются никем из испытуемых (пересечение негативных словарей):

- **Ядерные слова менее предметны (это видно как по соотношению слов по частям речи, так и по семантическим категориям),**
- **менее стилистически маркированы,**
- **более многозначны (а именно, средняя общая полисемия слов этого списка существенно выше - **4,0 значения на 1 слово**), чем слова всего тестового списка в целом (**1,98 значения на 1 слово**),**
- **а также все эти слова представлены в словнике Ожеговского словаря.**

Слова же списка слов, никем не употребляемых и даже никому из испытуемых не известных, в отличие от первого списка, относятся к крайней периферии словарного состава русского языка по их характеристикам. Они практически **неполисемичны, стилистически сильнее маркированы** (и маркированы преимущественно **специальной сферой употребления - более 17% значений этих слов**), а также **представлены в словаре Ожегова намного меньше**, чем, в среднем, все слова тестового списка.

4. Выводы.

Существующие в настоящее время системы лексического тестирования ориентированы на решение сугубо практических задач, преимущественно, для аттестации и сопровождения учебного процесса. Глобальные научные задачи, типа определения границ трёхслойной структуры словарного сознания общества ни в одном исследовании до настоящего времени не ставились. Главным в настоящем исследовании является то, что тестирование словарного знания индивидуумов производится по большому тестовому списку (ок. 1650 единиц), в максимальной степени отображающему различные сферы и слои русского языка, с дифференциацией при тестировании активного и пассивного знания (а также – незнания) лексических единиц испытуемыми, с учётом различных общеязыковых характеристик лексических единиц и их представленности в основных толковых словарях русского языка. Подобный комплекс задач в лингвистике ещё не рассматривался. Точно так же, как и не ставилась задача проецирования полученной по тестовому списку границы активного «ядра» (по пересечению активных индивидуальных словарей) на полный состав

лексики современного русского языка – с целью получения типичного портрета состава этого активного «ядра».

Реализуемый в настоящем проекте подход осуществляется с позиций **СИСТЕМНОЙ ЛЕКСИКОЛОГИИ**. Основным здесь является структурированное представление индивидуально-социальной онтологии языка и его словаря, позволяющее поставить вопрос о центральной роли в общении между носителями данного языка **ПЕРЕСЕЧЕНИЯ ИХ ИНДИВИДУАЛЬНЫХ АКТИВНЫХ СЛОВАРЕЙ**. На этой основе проводится последовательное соположение единиц тестового списка с системой языка – как по их представленности в толковых словарях разного типа, так и по их общеязыковым характеристикам. В том числе, лексические единицы характеризуются по денотативно-смысловым категориям их значений.

5. ПЕРСПЕКТИВЫ

Дальнейшее развитие этой серии экспериментов должно заключаться в уточнении и добавлении некоторых общеязыковых характеристик лексических единиц тестового списка (например, добавлении частотных характеристик по нескольким представительным частотным словарям русского языка), количественном и качественном расширении круга испытуемых (по возрасту, полу и уровню образования), привлечении средств математической статистики к углублённому анализу получаемых эмпирических данных, определении степени варьированности объёмов активных и пассивных словарей (и их пересечений и объединений) в зависимости от разных категорий испытуемых (по полу, возрасту, уровню образования). Одним из основных главных экспериментальных результатов должно стать статистическое определение **ВЕРОЯТНОСТНЫХ ГРАНИЦ ПЕРЕСЕЧЕНИЯ АКТИВНЫХ СЛОВАРЕЙ** для разных категорий испытуемых, что может стать основой для определения научно оправданных границ краткого и среднего, среднего и большого словарей современного русского языка, а также выработки объективно подобранных тестовых списков и словарей минимумов различного уровня при обучении русской лексике.

Вырабатываемая методика может быть приложена и к материалу других языков.

Литература

- БАС - Словарь современного русского литературного языка/ АН СССР. Ин-т рус. яз. М., Л., 1948 - 1965. Т. 1 - 17
- МАС - Словарь русского языка: В 4 т./ АН СССР. Ин-т рус. яз. М., 1957 - 1961. Т. 1 - 4
- МАС-2 - Словарь русского языка: в 4 тт. / Гл. ред. А.П. Евгеньева; АН СССР. Ин-т рус. яз. 2-е изд., испр., доп. М., 1981 - 1984. Т. 1 - 4
- Ожегов С.И. О трех типах толковых словарей современного русского языка // ВЯ, 1952, N 2.

- Ожегов С.И. Словарь русского языка / Под ред. Н.Ю. Шведовой. 9-е изд. - М., 1972, 848 с.
- Поддубный В. В., Поликарпов А. А. Вывод закона синхронного полисемического распределения языковых знаков на основе диссипативной стохастической динамической модели эволюции знаковых ансамблей // Синхронное и диахронное в сравнительно-историческом языкознании: Материалы VII Междунар. науч. конф. — М.: “Добросвет”, Изд-во “КДУ”, 2011. — С. 182–190.
- Поддубный В.В., Поликарпов А.А. Диссипативная стохастическая динамическая модель развития языковых знаков // КОМПЬЮТЕРНЫЕ ИССЛЕДОВАНИЯ И МОДЕЛИРОВАНИЕ 2011 Т. 2 № 1 С. 3–24.
- Поликарпов А.А. Элементы теоретической социолингвистики. -М.: Изд. МГУ, 1979. – 163 стр.
- Поликарпов А.А. Полисемия: системно-квантитативные аспекты // Квантитативная лингвистика и автоматический анализ текстов. Тарту, 1987 (Уч. зап. Тартуского гос. ун-та. Вып. 774).
- Поликарпов А.А. Системно-квантитативный подход в лингвистике // Научные школы в лингвистике. - Смоленск, 2007.
- А.А.Поликарпов. Универсальные, типологические, локально-смысловые и частнолингвистические факторы исторической сменяемости знаковых единиц // VIII Международная конференция по языкам Дальнего Востока, Юго-Восточной Азии и Западной Африки (Москва, 23-24 сентября 2009 г.): тезисы и доклады. -М.: КлючС, 2009. -С. 125-141.
- Поликарпов А.а. Системно-квантитативный анализ лексической полисемии китайского языка во взаимосвязи со структурными и стилистическими характеристиками слов //Вестник Смоленского государственного университета, № 4 (8). – Смоленск: Изд-во Смоленского университета, 2009а. -Сс. 133-152.
- Поликарпов, А.А. Закономерности активного и пассивного знания и незнания лексической системы русского языка московскими студентами // «Образ России и россиянина в словаре и дискурсе: когнитивный анализ. Тезисы докладов и сообщений Международной научной конференции, посвящённой юбилею заслуженного деятеля науки РФ, доктора филологических языка, профессора Л.Г.Бабенко 29.09 – 1.19.2011 г. Екатеринбург, Россия. – Екатеринбург: Издательство Уральского университета, 2011.
- Поликарпов А.А., Крюкова О.С. О системном соотношении краткого и среднего толковых словарей русского языка // Квантитативная лингвистика и автоматический анализ текстов. Тарту, 1989 (Уч. зап. Тартуского ун-та. Вып. 872).
- Поликарпов А.А., Курлов А.А. Стилистика, семантика, грамматика. Опыт анализа системных взаимосвязей (По данным толкового словаря) // ВЯ, 1994, N 1. - С. 62-75
- Поликарпова А.О., Поликарпов А.А. Опыт изучения уровня и характера индивидуального знания русской лексики // Квантитативные аспекты

- системной организации текста: Материалы Межвуз. семинара. - Тбилиси, 1987.
- Рогожникова Р.П. Сводный словарь современной русской лексики. В 2-х тт. - М.,1991.
- Савчук Л.О. Системные соотношения словаря общества и словаря индивида (с точки зрения активного и пассивного словарного запаса). Автореф. канд. дисс. - М.,1997.
- Словарь-тезаурус синонимов русской речи. (Под ред. Бабенко Л.Г.). -М.: АСТПРЕСС КНИГА, 2007.
- В.П. Тимофеев. Диалектный словарь личности. (Шадринск, 1971.
- Толковый словарь русского языка / Под ред. Д.Н. Ушакова. М., 1935 - 1940
- Polikarpov, A.A. 1993. On the Model of Word Life Cycle // Koehler, R., Rieger, B. (eds.) Contributions to Quantitative Linguistics.-Dordrecht: Kluwer, London,Boston, 1993. - Pp. 53-66.
- Polikarpov, A.A. 2004. Cognitive Mechanisms for Emergence of some Regularities of Human Language Evolution // Text Processing and Cognitive Technologies. № 9,2004 (The VII-th International Conference "Cognitive Modeling in Linguistics.Proceedings"). – Moscow – Varna, 2004. - Pp. 10-20.
- Polikarpov A.A. 2006. Towards the Foundations of Menzerath's Law (On the Functional Dependence of Affix Length on their Positional Number within Words) // Peter Grzybek (ed.). Contributions to the Science of Text and Language. Word Length Studies and Related Issues. - Dordrecht: Springer, 2006. –Pp. 215-240.
- Polikarpov A.A. Asymmetric dualism of a language sign: on the mechanism of establishing harmony between polysemy and synonymy in natural language signs //Abstracts from Qualico-2009 (Quantitative Linguistics Conference: «Text and Language: Structures, Functions, Interrelations», Graz, Austria, September 17-20 2009). – Graz, Graz University, 2009.