

УДК 81-33

КОРПУС РУССКИХ ПИСЬМЕННЫХ ТЕКСТОВ ЛИЦ ПОЗДНЕГО ВОЗРАСТА: МЕТОДИКА СБОРА МАТЕРИАЛА

ЛАНТЮХОВА Наталья Николаевна,аспирант кафедры русского языка, современной русской и зарубежной литературы,
Воронежский государственный педагогический университет

АННОТАЦИЯ. В статье рассмотрены некоторые особенности современных отечественных корпусов текстов. Подробно описана методика создания специального корпуса текстов лиц позднего возраста *RusOld* с помощью разработанной анкеты, обоснован состав анкеты. Формируемый корпус предназначен в первую очередь для исследования синтаксических особенностей речи лиц старшей возрастной группы (старше 60 лет). По нашим данным, подобное исследование на материале русского языка проводится впервые.

КЛЮЧЕВЫЕ СЛОВА: корпус текстов, старение, пожилой возраст, метатарзметка, экстралингвистические данные.

THE CORPUS OF RUSSIAN WRITTEN TEXTS OF PEOPLE ADVANCED IN AGE: METHODOLOGY OF MATERIAL COLLECTION

LANTYUKHOVA N. N.,Post-Graduate Student of the Department of Russian Language, Modern Russian and Foreign Literature,
Voronezh State Pedagogical University

ABSTRACT. The article considers some features of modern domestic corpus of texts. The methodology of creating a special corpus of advanced age persons' texts (*RusOld*) based on the material of the Russian language with the help of the questionnaire is described in detail, the questionnaire content is substantiated. The formed corpus is designed to study the syntactic features of speech of the older age group (over 60 years old). According to our data, such a study based on the material of the Russian language is carried out for the first time.

KEY WORDS: corpus, aging, advanced age, metatagging, extralinguistic data.

Корпусные исследования в современной лингвистике занимают значительное место, поскольку использование корпусов, то есть «массивов текстов, собранных в единую систему по определенным признакам (языку, жанру, времени создания текста, автору и т.п.) и снабженных поисковой системой» [11, с. 99], позволяет выявлять и статистически подтверждать различные языковые явления, проследить изменения в языке [6, с. 387-388].

В настоящее время выделяют различные виды корпусов: относящиеся ко всему национальному языку (полные) и относящиеся к какому-либо подязыку (специальные), созданные специально для наблюдения и изучения какого-либо явления (корпус пословиц, корпус политических метафор, корпус газетной речи и т.д.), статические, включающие тексты какого-либо небольшого временного промежутка, и динамические, относящиеся к большому периоду, корпусы параллельных текстов, состоящие из множества текстов-оригиналов и их переводов на другие языки, эталонные корпуса, отражающие языковую норму, интернет-корпуса, содержащие подборку текстов из интернета, отобранных по каким-то критериям, и т.д.

Несмотря на большое количество корпусов на материале русского языка, в том числе и очень объемных (например, Национальный корпус русского языка в настоящее время насчитывает более 600 млн словоупотреблений, а Генеральный интернет-корпус русского языка около 20 миллиардов словоупотреблений), собранные коллекции текстов не

всегда отвечают нуждам исследователей, что вызывает появление различных специальных корпусов.

Например, в большинстве корпусов, размещенных в открытом доступе, крайне скупо представлена информация об авторах текстов (как правило, указаны только пол и возраст). Между тем информация о личности автора значительно расширяет возможности использования корпуса, позволяя не только анализировать какое-либо языковое явление, но и исследовать особенности функционирования этого явления в определенных социальных группах, изучать зависимости и вычислять корреляции между этим явлением и определенными характеристиками личности авторов. Установлено (см., например, [15-17]), что экстралингвистические факторы оказывают существенное влияние на речевое поведение человека, отражаясь не только на его содержании, но и структурных и качественных характеристиках текста.

Кроме того, использование метатарзметки, предполагающей данные о социальных и психофизиологических характеристиках авторов текстов, выводит языковые корпуса за рамки собственно лингвистики, делая их объектом исследования также историков, социологов, психологов, литературоведов и др.

Еще один отмеченный нами аспект: в корпуса (в первую очередь полные корпуса, относящиеся ко всему национальному языку) часто включаются тексты, оценить оригинальность которых довольно сложно. Это, например, тексты литературных произведений, статьи из газет и журналов, то есть все те источники, над которыми, кроме самого автора, работает штат издательских работников (корректоры, редакторы, литературные обработчики). В этом

случае, даже если известна вся информация об авторе, трудно объективно оценить его личный вклад в создание текста. А значит, эти тексты не могут быть использованы, например, в работах, связанных с исследованиями функционирования языка в определенных социальных группах, даже если сам автор принадлежит к этой группе.

Кроме того, основная часть текстов в корпусах (особенно «лабораторных» корпусах, создаваемых респондентами по заданию исследователя) создается, как правило, наиболее активной частью социума – лицами в возрасте 20-40 лет. Старшая возрастная группа (лица пожилого и старческого возраста) оказываются представленными в корпусах минимально.

Все это привело нас к необходимости создания для целей нашего исследования - изучения особенностей речи лиц позднего возраста (старше 60 лет) - нового, специального корпуса текстов *RusOld*. О существовании подобных специальных корпусов на материале русского языка нам неизвестно.

Создаваемый нами корпус *RusOld* формируется в рамках работ Регионального центра русского языка Воронежского государственного педагогического университета (<http://ruscentre.nsknet.ru>) и с учетом достижений Лаборатории корпусной социолингвистики (<http://rusprofilnglab.ru>).

В корпусе *RusOld* представлены образцы естественной письменной речи, снабженные развернутой информацией об авторах.

Сбор материала для исследования проводился с помощью специально разработанной анкеты, состоящей из ряда блоков. Время на заполнение анкеты не ограничивалось, за исключением интеллектуального тестирования, на которое отводилось 20 минут. В среднем заполнение всей анкеты занимало у испытуемых около полутора часов.

В связи с часто наблюдаемым в позднем возрасте снижением зрения и могущим возникнуть в связи с этим чувством неуверенности, дискомфорта [3, с. 85-99; 8, с. 719-759; 14, с. 44-45] анкета была набрана крупным шрифтом со значительными интервалами. Поскольку психологическое состояние может существенно повлиять на результаты испытания.

1 блок предназначен для выявления общей информации об испытуемых, которая может коррелировать с грамматическими параметрами текста:

- пол;
- возраст. В исследовании принимают участие лица в возрасте 60 и более лет, верхняя возрастная планка не ограничивалась;
- образование (также указывался профиль, если образование высшее).

По данным целого ряда исследователей [2, с. 137-246; 10, с. 112-116; 1, с. 88-94; 14, с. 32-33; 13, с. 59-60 и др.], образование является мощным механизмом, противостоящим инволюции, способным буквально затормозить старение мозга. Особенно значительное влияние фактор образования оказывает на сохранность вербальных способностей человека. Так, например, интеллектуальное тестирование пожилых с помощью тестов Векслера показывает, что успешность выполнения трех вербальных тестов – словарный запас, информированность и общая осведомленность – напрямую связано с уровнем образования: если последний достаточно высок, можно не наблюдать никакого спада до глубокой старости, а иногда даже улучшение показателей до поздней зрелости [1, с. 61-86];

- следующий блок вопросов относится к трудовой деятельности респондента (ваш трудовой

стаж; ваша профессия; работаете ли вы в настоящее время; если нет, сколько лет вы не работаете; имее ли вы хобби (увлечение), которому систематически уделяете много времени).

Психологи [3, с. 85-99; 8, с. 719-759; 14, с. 44-45; 2, с. 188; 9, с. 141] отмечают такое частое явление, как деформация личности, связанная с прекращением профессиональной трудовой деятельности. Даже при так называемом нормальном старении, не отягощенном психическими и физическими заболеваниями, влияющими на интеллект, многие используют свое право ухода на пенсию, хотя физическое и интеллектуальное состояние вполне позволяет работать дальше. В то же время люди не имеют каких-либо увлечений, способных заменить эту работу. Это приводит к резкому сужению объема поступающей информации, когда отпадает необходимость вникать в рабочие вопросы, осваивать новые навыки, изучать новые данные. Функции, связанные с работой, которые ранее постоянно тренировались, начинают ослабевать, происходит по выражению Б.Г. Ананьева деформация личности [2, с. 188].

В этом контексте наличие хобби у пожилых становится важным аспектом, влияющим на интеллектуальные, в том числе и вербальные, способности личности. Например, увлечение цветоводством заставляет человека не только проявлять хотя бы минимальную физическую активность, связанную с уходом за растениями, но также постоянно искать и осваивать информацию о своем хобби (узнавать о новых растениях, методах ухода за ними), общаться с единомышленниками;

- блок вопросов, характеризующий социальную адаптацию респондента, «встроенность» его в общественную жизнь, интерес к окружающему миру, к новой информации (интересуетесь ли вы общественной жизнью, политической ситуацией в стране и мире; вы ходите на выборы; вы умеете работать с компьютером; вы зарегистрированы в каких-либо социальных сетях; сколько книг вы прочли за последний год; как вы часто общаетесь с друзьями, родственниками).

Этот блок вопросов логически связан с предыдущим: отказ от некоторых социальных функций, сужение интересов, уменьшение количества поступающей информации, вызванные выходом на пенсию, часто ведут к добровольной социальной изоляции. Также такая изоляция может быть и вынужденной, связанной с умиранием ровесников, физическими заболеваниями, связанной с этим депрессией. Все это приводит к постепенному сужению когнитивных функций [9, с. 141; 3, с. 85-99; 2, с. 188; 8, с. 719-759; 14, с. 44-45].

Отдельно отметим вопросы, связанные с умением работать с компьютером: поскольку персональные компьютеры в нашей стране стали массово появляться только во 2-й половине 90-х годов, многие респонденты учились работать на них уже в довольно зрелом возрасте, то есть получали дополнительные образовательные навыки, тренировку мелкой моторики.

2 блок направлен на выявление пожилых, страдающих возрастной деменцией. В данной работе эта категория лиц не рассматривалась, а их анкеты не учитывались.

Тестирование проводилось с помощью теста рисования часов [7]. Человека просят самостоятельно нарисовать круглый циферблат часов и с цифрами и стрелками, которые бы указывали заданное ис-

следователем время. При этом используется чистая белая бумага, без линий и клеток.

Данная методика является одной из самых простых и в тоже время высокочувствительных.

3 блок направлен на выявление лиц, страдающих депрессией и находящихся в тревожном состоянии. Оценка состояния проводилась с помощью «Госпитальной шкалы тревоги и депрессии» [4]. Анкеты людей, у которых была выявлена депрессия или тревожность, обрабатывались наравне с другими.

Депрессия и тревожность приводят к снижению концентрации внимания, нарушению памяти, в целом понижают уровень когнитивной деятельности и могут опосредованно значительно влиять на грамматические особенности текста. Часто этот страх, неудовлетворенность прожитой жизнью и нынешним положением. Человек «закрывается», отдаляясь от людей (друзей, родственников, соседей) и от мира в целом: не интересуется новостями, жизнью близких – мир сужается. Человек начинает жить в основном прошлом, новое его не интересует и воспринимается с трудом [12; 8, с. 145; 8, с. 719-759 и др.].

4 блок. Участникам исследования предлагалось написать сочинение по одной из предложенных тем. Никаких иных ограничений не ставилось. Приветствовался большой объем текста, допускались правки. Темы подобраны таким образом, чтобы продемонстрировать связь человека с окружающим миром во всем его многообразии, его психологическое состояние, вызвать эмоции, заставить рассуждать.

На выбор предлагались следующие темы: «Мои дети», «Человек и его дело (моя профессия)», «Самый счастливый день в моей жизни (свадьба, рождение детей, другое)», «Самый печальный день моей жизни», «Мой самый хулиганский поступок», «Каким было мое детство», «Мое хобби», «Друзья», «Мой враг», «Что я планирую сделать в следующем году», «Современная Россия моими глазами», «Путешествия в другие города».

5 блок – интеллектуальное тестирование.

В большинстве случаев оценка психологических и интеллектуальных изменений в старости проводится с помощью различных тестов, что вызывает обоснованную критику ученых, не согласных с формально-логическим подходом к мышлению [8, с. 719-759; 1, с. 90; 3, с. 85-99 и др.]. При этом игнорируются глубинные пласты мышления, их особая логика, которая актуализируется у пожилых людей. Изучая ход выполнения тестов интеллекта в разные периоды жизни, психологи показывают, что индивиды на этапе поздней жизни относятся к тестовым заданиям как к жизненным проблемам, они персонализируют их, дают им несколько интерпретаций, видят возможность разных ответов. Они опираются на интуитивные процессы, прибегают к метафорам, выходят за пределы данной им информации. Ответ, построенный лишь на формально-логическом мышлении, представляется им бедным и односторонним.

Еще одним возражением против интеллектуального тестирования является то, что скорость ответа в тестах регламентирована. Учитывая доказанное снижение времени реакции в старости, очевидно, что тесты всегда будут показывать интеллектуальную недостаточность лиц пожилого и старческого возраста по сравнению с молодыми. В то же время низкая скорость может быть результатом того, что пожилые люди больше, чем молодые, ценят аккуратность и точность. При тестировании пожилые люди делают меньше предположений и стараются

точнее отвечать на каждый вопрос. Опять же возникает вопрос о том, что именно измеряют тесты – скорость реакции или собственно интеллект.

Другой значимый момент – пожилые люди часто плохо знакомы с некоторыми заданиями, использованными в тестовых ситуациях. Если, например, современные молодые люди постоянно сталкиваются с тестированиями, начиная со школы, то многие пожилые люди просто теряются, впервые столкнувшись с такой методикой. Поэтому сравнение тестов пожилых людей и молодых иногда приводит к неверным выводам и имеет мало общего с реальным положением дел.

Нельзя также не принять во внимание выводы некоторых исследователей о том, что результаты лабораторных экспериментов расходятся с данными, полученными при изучении пожилых людей в их повседневной жизни. В лабораторных экспериментах часто исследуются когнитивные процессы, нерелевантные реальной жизни людей.

Несмотря на эти возражения, тестирование по-прежнему остается наиболее популярным методом массовой оценки интеллекта, в том числе и у лиц позднего возраста.

В нашем случае тестирование проводилось с помощью Краткого ориентировочного теста (КОТ) Э.Ф. Вандерлика в адаптации В.Н. Бузины. Данный тест был выбран с учетом, во-первых, его возможностей в определении не только интегрального показателя общих способностей человека, но и отдельных характеристик личности (математические способности, пространственные, вербальные и др.), а во-вторых – с учетом удобства его использования в силу достаточной краткости, что представляется значимым на фоне большого размера анкеты в целом.

Таким образом, предлагаемая респондентам анкета построена так, чтобы, помимо собственно текстов, получить максимальное количество информации о социальных, психологических и интеллектуальных характеристиках авторов, что позволит исследовать влияние этих характеристик на собственно лингвистические параметры порождаемого текста.

В настоящее время с помощью анкеты собрано более 200 текстов, и корпус постоянно пополняется. Тексты собираются на бумажных носителях и далее переводятся в электронный формат для дальнейшего анализа, что связано преимущественно с возрастом респондентов, многим из которых сложно работать с компьютером долгое время.

Сбором материала занимались 4 волонтера, двое из них принадлежали к той же возрастной группе, что и опрашиваемые (67 и 69 лет, женщины), двое среднего возраста (35 – мужчина и 37 – женщина). В анкетировании принимали участие как лично знакомые волонтерам люди (≈15 %), так и незнакомые. Вознаграждение за участие в исследовании не предлагалось и не выплачивалось.

Работа с каждым респондентом велась индивидуально, все участники исследования (в том числе и те, кто затем отказался участвовать в опросе) активно интересовались целью работы, задавали уточняющие вопросы. Отметим, что наиболее продуктивно работали именно волонтеры старшей возрастной группы, они легко вступали в контакт с ровесниками, практически не получали отказов в сотрудничестве (менее 5%). В то же время более молодые интервьюеры получали отказ примерно в 30% случаев. Причины отказа заполнять анкету также назывались разные в зависимости от возраста исследователей: более молодым интервьюерам отказ участвовать в работе объяснялся преимущественно

отсутствием свободного времени, тогда как представителям своей возрастной группы респонденты признавались в страхе не справиться с заданиями, страхе несостоятельности («Не хочу выглядеть глупой», «Не умею выполнять тесты», «Я и в школе не умел писать сочинения», «У меня нет мыслей, надо мной будут смеяться», «У меня плохо с математикой» и подобные). Несмотря на то, что в анкете не указывалась фамилия опрашиваемого (и в большинстве случаев исследователи ее просто не знали), а волонтеры подробно объясняли условия использования материала, страх несостоятельности, отсутствие уверенности в себе и своих силах оказались самым значительным препятствием в сборе материала.

В то же время респонденты, согласившиеся участвовать в исследовании, относились к работе, как правило, очень ответственно: отвечая на вопросы, многие оставляли комментарии на полях, поясняя свой ответ, часто устно комментировали волонтерам ход решения задач интеллектуального тестирования.

Вообще наибольший интерес у респондентов в большинстве случаев вызвали КОТ и тест рисования часов, их выполняли охотно, интересовались результатами. Цель выполнения теста рисования часов (выявление лиц, страдающих возрастной деменцией) в анкете не указывалась по этическим соображениям, и это задание воспринималось опрашиваемыми как еще одна интеллектуальная задача.

В отношении 5-го блока – сочинения – наблюдалось две прямо противоположных реакции. Примерно для трети респондентов это оказалось самым сложным заданием: они быстро и довольно легко справились с остальными четырьмя блоками, но сочинение либо не могли написать вовсе (сдавали анкету без заполнения этого блока, хотя волонтеры объясняли необходимость заполнения всей анкеты), либо же ограничивались двумя-тремя предложениями. В то время как для остальных сочинение стало наиболее интересным заданием, респонденты писали значительные по объему тексты, вносили много правок, переписывали текст, чтобы точнее сформулировать мысль.

Работа по сбору материала с помощью разработанной анкеты продолжается.

В то же время корпус был дополнен также материалами из блогов пожилых интернет-пользователей (также старше 60 лет), размещенными преимущественно на базе «Живого журнала» (<https://www.livejournal.com>), а также из личных блогов. Эта часть коллекции была добавлена для сравнения, выявления возможных особенностей, связанных со спецификой ведения сетевых дневников.

Материал отбирался исходя из следующих условий:

- 1) год рождения пользователя указан в личных данных (записи, сделанные до наступления 60-летия, не учитывались);
- 2) сведения о блогере заполнены максимально подробно, в целом соответствуют блоку 1 анкеты;
- 3) блог не является тематическим (не учитывались блоги, посвященные, например, кулинарии, охоте, рукоделию, рыболовству, садоводству и т.д.);
- 4) записи в дневнике авторские, не являются репостом чужих материалов;
- 5) записи в целом соответствуют темам, указанным в блоке 5, несут нарративный характер.

Одна запись учитывалась как один текст. Отметим, что блоги пожилых пользователей, как правило, ведутся неактивно (по сравнению с блогами более молодых авторов), записи появляются с большими перерывами и их немного. В настоящее время собрано более 100 текстов. Сбор материала продолжается.

Корпус в дальнейшем будет выложен в интернете и будет доступен всем исследователям.

Создаваемый корпус за счет метаразметки и строгой ориентированности на определенную социальную группу, по нашему мнению, может быть интересен не только лингвистам, но также найдет применение в социологии, психологии, геронтологии, культурологии, истории и других сферах. Материалы названного корпуса могут быть положены в основу работ по лексикографическому портретированию [5] представителей старшей возрастной группы носителей русского языка, что представляется чрезвычайно важным для многих современных междисциплинарных исследований.

СПИСОК ЛИТЕРАТУРЫ:

1. Александрова, М.Д. Проблемы социальной и психологической геронтологии [Текст] / М.Д. Александрова. – Л.: Изд-во Ленинград. ун-та, 1974. – 136 с.
2. Ананьев, Б.Г. О проблемах современного человекознания [Текст] / Б. Г. Ананьев. – 2-е изд. – СПб: Питер, 2001. – 272 с.
3. Анцыферова, Л.И. Психология старости: особенности развития личности в период поздней зрелости [Текст] / Л.И. Анцыферова // Вопросы психологии. – 2001. – № 3. – С. 85-99.
4. Воробьев, П.А. Депрессия в пожилом возрасте [Текст] / П.А. Воробьев, А.В. Власов // Клиническая геронтология. – 2007. – № 3. – С. 22-28.
5. Загоровская, О.В. Речевое портретирование социальной группы как задача современной лексикографии [Текст] / О.В. Загоровская // Вестник Воронежского государственного университета. Серия: Лингвистика и межкультурная коммуникация. – 2013. – № 2. – С. 174-177.
6. Загоровская, О.В. Электронный корпус студенческих эссе на русском языке и его возможности для современных гуманитарных исследований [Текст] / О.В. Загоровская, Т.А. Литвинова // Мир науки, культуры, образования. – 2012. – № 3 (34). – С. 387-389.
7. Захаров, В.В. Когнитивные расстройства в пожилом и старческом возрасте [Текст] / В.В. Захаров, Н.Н. Яхно. – М., 2005. – 71 с.
8. Крайг, Г. Психология развития [Текст] / Г. Крайг, Д. Бокум. – 9-е изд. – СПб: Питер, 2005. – 940 с.
9. Мухордова, О.Е. Особенности познавательной сферы женщин пожилого и старческого возраста [Текст] / О.Е. Мухордова // Вестник Удмуртского университета. – 2009. – Вып. 3-2. – С. 139-146.
10. Носко, И.В. Психология развития и возрастная психология [Текст] / И. В. Носко. – Владивосток: Дальневосточный гос. университет, 2003. – 131 с.

11. Сысоев, П.В. Лингвистический корпус в методике обучения иностранным языкам [Текст] / П.В. Сысоев // Язык и культура. – 2010. – № 1 (9). – С. 99-111.
12. Хухлаева, О. В. Психология развития: молодость, зрелость, старость [Текст] / О.В. Хухлаева. – М.: ИЦ «Академия», 2002. – 208 с.
13. Штернберг, Э.Я. Геронтологическая психиатрия [Текст] / Э.Я. Штернберг. – М.: Медицина, 1977. – 216 с.
14. Яцемирская, Р.С. Психопатология пожилого и старческого возраста : курс лекций [Текст] / Р.С. Яцемирская. – М.: РГСУ, 2001. – 212 с.
15. Cheung, H. Competing complexity metrics and adults' production of complex sentences / H. Cheung, S. Kemper // Applied Psycholinguistics. – 1992. – № 13. – P. 53-76.
16. Kemper, S. Life-span changes in syntactic complexity / S. Kemper // Journal of Gerontology. – 1987. – № 42 (3). – P. 323-328.
17. Kemper, S. The structure of verbal abilities in young and older adults / S. Kemper, A. Sumner // Psychology and Aging. – 2001. – № 16 (2). – P. 312-322.