ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

на тему:

**Дистантные связи в предикатных группах русской разговорной речи (на материале корпуса «Один речевой день»)**

основная образовательная программа магистратуры по направлению подготовки 45.04.02 «Лингвистика»

Исполнитель:

Обучающийся 2 курса магистратуры

Образовательной программы

«Русский язык»

очной формы обучения

Наумова Анфиса Сергеевна

Научный руководитель:

к. ф. н., доц. Т. Ю. Шерстинова

Рецензент:

к. ф. н., доц. П. М. Эйсмонт

Санкт-Петербург

2018

# ОГЛАВЛЕНИЕ

[ОГЛАВЛЕНИЕ 2](#_Toc514356571)

[ВВЕДЕНИЕ 4](#_Toc514356572)

[ГЛАВА 1. РУССКАЯ РАЗГОВОРНАЯ РЕЧЬ КАК ОБЪЕКТ ЛИНГВИСТИЧЕСКОГО ИССЛЕДОВАНИЯ 9](#_Toc514356573)

[1.1. Дихотомия «язык — речь» 9](#_Toc514356574)

[1.2. Особенности разговорной речи 11](#_Toc514356575)

[1.3. Корпусные методы в исследованиях устной речи 14](#_Toc514356576)

[1.3.1. Обзор наиболее известных корпусов для русского языка 16](#_Toc514356577)

[1.3.2. Звуковой корпус русского языка 18](#_Toc514356578)

[1.4. Социолингвистический аспект в исследованиях устной речи (и синтаксиса разговорного языка) 21](#_Toc514356579)

[1.4.1. Пол говорящего как социолингвистический фактор 25](#_Toc514356580)

[1.4.2. Возраст говорящего как социолингвистический фактор 28](#_Toc514356581)

[1.4.3. Профессия говорящего как социолингвистический фактор 31](#_Toc514356582)

[ГЛАВА 2. СИНТАКСИЧЕСКИЕ ОСОБЕННОСТИ РУССКОЙ РАЗГОВОРНОЙ РЕЧИ 34](#_Toc514356583)

[2.1. Синтаксис устной и письменной речи 34](#_Toc514356584)

[2.2. Синтаксические исследования русской разговорной речи 34](#_Toc514356585)

[2.3. Подходы к анализу синтаксических структур 38](#_Toc514356589)

[2.4. Симметрия дистантных связей в устной и письменной речи 39](#_Toc514356590)

[ГЛАВА 3. ИССЛЕДОВАНИЕ ДИСТАНТНЫХ СВЯЗЕЙ В ПРЕДИКАТНЫХ ГРУППАХ РУССКОЙ РАЗГОВОРНОЙ РЕЧИ 41](#_Toc514356591)

[3.1. Материал и методика 41](#_Toc514356593)

[3.1.1. Линейный синтаксис 43](#_Toc514356594)

[3.1.2. Понятия предикатной группы и синтаксического элемента 44](#_Toc514356595)

[3.1.3. Синтаксические категории 45](#_Toc514356596)

[3.1.4. Методы формального подсчёта 46](#_Toc514356597)

[3.2. Результаты 47](#_Toc514356598)

[3.2.1. Дистантные связи в устной и письменной речи 47](#_Toc514356604)

[3.2.2. Дистантные связи в речи различных соцгрупп 49](#_Toc514356605)

[3.3. Выводы 56](#_Toc514356606)

[ГЛАВА 4. POS-АНАЛИЗ ПРЕДИКАТНЫХ ГРУПП 58](#_Toc514356607)

[4.1. Частота встречаемости различных типов ПГ 58](#_Toc514356609)

[4.2. Типичные структуры ПГ 62](#_Toc514356610)

[4.3. Выводы 72](#_Toc514356611)

[ЗАКЛЮЧЕНИЕ 75](#_Toc514356612)

[СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ 77](#_Toc514356613)

[СПИСОК СОКРАЩЕНИЙ, ПРИНЯТЫХ В РАБОТЕ 93](#_Toc514356614)

[ПРИЛОЖЕНИЯ 94](#_Toc514356615)

[Приложение. Таблицы 94](#_Toc514356616)

# ВВЕДЕНИЕ

Устная спонтанная речь стала предметом исследований специалистов в области языкознания сравнительно недавно. В оппозиции устная vs. письменная речь последняя долгие годы занимала главенствующее положение в лингвистических исследованиях. Что касается устной речи, она становилась объектом научного анализа существенно реже и, как правило, с позиций диалектологии. Однако во второй половине прошлого века средства фиксации речи стали более распространёнными и легкодоступными, что обусловило рост научного интереса к ней. Успешными направлениями в исследовании диалогической речи являются теория речевых актов [Остин, 1986; Серль, 1986] и конверсационный анализ [Sacks, Schegloff, 1973; Sacks, Schegloff, Jefferson, 1974]. Монологическая же речь привлекла внимание исследователей немного позднее; здесь в первую очередь следует отметить работы социолингвистов — Б. Н. Головина [Головин, 1969], Л. П. Крысина [Крысин, 1989], Т. И. Ерофеевой [Ерофеева Т.И., 1991] и др.

Любыми формами устной речи позволяет заниматься корпусная лингвистика — «область лингвистики, связанная с созданием и совершенствованием корпусов текстов, а также с их применением в качестве инструмента лингвистического исследования» [Яскевич, 2009]. Развитие корпусной лингвистики послужило разностороннему анализу устной речи и её особенностей на всех уровнях языка, начиная с фонетического и заканчивая дискурсивным. Так, например, на синтаксическом уровне выделяют единицы, которые могли бы послужить аналогом предложения в письменной форме языка, например, высказывания [Бахтин, 1986, с. 258], клаузы [Тестелец, 2001, с. 229–268], элементарные дискурсивные единицы (ЭДЕ) [Кибрик, Подлесская, 2009] и прочие.

В настоящее время одним из активно разрабатываемых корпусов русского языка является корпус «Один речевой день» [Asinovsky et al., 2009; Звуковой корпус …, 2013; 2014; 2015]. Особенностью данного ресурса является то, что он содержит звукозаписи повседневной русской речи, выполненные в условиях естественной коммуникации. Научный коллектив, работающий с данным корпусом, исследует устную спонтанную речь в различных аспектах. Однако не до конца выясненным остаётся вопрос организации речи на синтаксическом уровне.

В связи с вышесказанным ***объектом*** настоящего исследования стала русская разговорная речь, представленная в речевом корпусе русского языка «Один речевой день» (ОРД корпус). ***Предметом*** исследования были выбраны предикатные группы (ПГ). Под предикатной группой в данной работе понимается предикат и его окружение. В отличие от традиционно используемого термина «предикативная группа» в предикатную в настоящем исследовании включаются не только зависимые от предиката единицы, но и формально не зависимые от него элементы, относящиеся однако к цепочке слов ПГ (например, частицы, междометья и др.). В частности, внимание было обращено на дистантные связи в предикатных группах. Под дистантными связями в настоящем исследовании понимаются левая и правая ширина предикатных групп (их дистантность).

***Целью*** данной квалификационной работы является описание и анализ синтаксиса повседневной русской речи с позиции дистантных связей в предикатных группах.

Цель, поставленная в настоящем исследовании, обусловила следующие конкретные ***задачи***, требующие решения:

1. изучить научную литературу, касающуюся синтаксиса русской устной речи и его единиц;
2. рассмотреть линейную удалённость зависимых синтаксических элементов от главного;
3. определить средний объём предикатной группы в устной речи, а также её правую и левую дистантность;
4. сопоставить полученные данные с количественными характеристиками высказываний в письменной речи;
5. проанализировать влияние гендерных и возрастных факторов на степень и характер дистанцирования элементов в рассматриваемых синтаксических группах.

***Материалом*** для исследования послужило 830 предикатных групп.

Материал анализировался с помощью традиционных методов лингвистического анализа (описательного, сопоставительного и др.), при формировании количественных подсчетов использовались *статистические* *методы анализа данных*.

Основная ***гипотеза***, на проверку которой ориентирована данная работа, такова: разговорная русская речь имеет тенденцию к левосторонности, т. е. в устной речи в отличие от письменной левая дистантность предикатных групп количественно преобладает над правой. Кроме того, в работе предпринимается попытка определить, существует ли зависимость между синтаксическим уровнем языка (в частности, дистантными связями предикатных групп) и социальными характеристиками говорящего.

***Актуальность*** настоящей работы заключается прежде всего в её исследовательской направленности — изучение спонтанной бытовой устной речи в последние годы служит предметом активного научного интереса в различных областях лингвистики. Данная работа является вкладом в создание грамматики речи, о необходимом создании которой в научном сообществе говорится уже полвека [Мартинэ, 1963, с. 389]. Также исследование затрагивает социальный аспект анализа речи, что соответствует растущему интересу к междисциплинарным научным изысканиям. Кроме того, исследование проведено в рамках корпусного подхода, одного из ведущих в современной лингвистике, а полученные результаты являются актуальными для дальнейших исследований синтаксических особенностей разговорной речи в рамках прикладной лингвистики.

***Научная новизна*** исследования обусловлена тем фактом, что впервые проводится многоаспектный анализ дистантных связей синтаксических групп на материале корпуса звукозаписей русской разговорной речи, получены статистические данные о реализации предикатных групп в речи представителей разных социальных групп и проведён их сопоставительный анализ.

В ***теоретическом плане*** работа вносит вклад в описание синтаксиса русской устной речи. Настоящее исследование включает комплексное описание количественных характеристик синтаксических единиц, выявленных на устном материале. Эти данные необходимы для построения грамматики разговорной речи, для документального подтверждения отличий устной речи от письменной на синтаксическом уровне. Кроме того, полученные данные имеют значение и для социолингвистических исследований повседневной русской речи.

***Практическое значение*** исследования заключается в систематизированном описании типовых синтаксических единиц устной речи. Эти данные могут использоваться в многочисленных практических приложениях, связанных с построением синтаксической модели языка повседневного общения (в частности, для разработки многофункциональных диалоговых систем при общении человека с автоматизированными устройствами на естественном языке, для систем синтеза и распознавания повседневной речи и др.).

***Структура работы*** отражает ход исследования и включает введение, 2 теоретических и 2 практических главы, заключение, список использованной литературы, список принятых в работе сокращений и приложение.

***Апробация работы*** была проведена в ходе обсуждения некоторых положений и результатов настоящего исследования на научных конференциях:

1. ХX Международная конференция студентов-филологов (17‑21 апреля 2017 г., Санкт-Петербург);
2. Международная конференция молодых филологов (28-30 апреля 2017 г., Тарту, Эстония);
3. IV Студенческая конференция Института лингвистики (12–13 апреля 2018 г., Москва);
4. XXI Международная конференция студентов-филологов (16-20 апреля 2018 г., Санкт-Петербург).

Основные положения настоящего исследования отражены в ***публикациях*** автора:

1. Количественные характеристики устных «предложений» в повседневной речи в социологическом аспекте // Русская филология. 29. — Тарту, 2017. — С. 461–473;
2. Далеко ли от «ну» до «ладно»? // XX Открытая конференция студентов-филологов СПбГУ, Санкт-Петербург, 17–21 апреля 2017 г.: Тезисы докладов. — СПб.: Филологический факультет СПбГУ, 2017. — C. 22.
3. Структура предикатных групп в устной спонтанной речи // XXI Открытая конференция студентов-филологов СПбГУ, Санкт-Петербург, 16–20 апреля 2017 г.: Тезисы докладов. — СПб.: Филологический факультет СПбГУ, 2018. — C. 45–46.
4. Далеко ли от «ну» до «ладно»? // Материалы XX Открытой конференции студентов-филологов. Санкт-Петербург, 17–21 апреля 2017 года. — СПб.: Издательство СПбГУ, 2017. — В печати.

# ГЛАВА 1. РУССКАЯ РАЗГОВОРНАЯ РЕЧЬ КАК ОБЪЕКТ ЛИНГВИСТИЧЕСКОГО ИССЛЕДОВАНИЯ

## Дихотомия «язык — речь»

Лингвистика рассматривает язык и речь как различные объекты исследования, начиная с работ Ф. де Соссюра [Соссюр, 1999; 2001]. Однако, несмотря на это, устная речь до сих пор иногда представляется менее престижной формой, чем письменная. Такой подход несправедлив: «речь <…> необходима для того, чтобы сложился язык; исторически факт речи всегда предшествует языку» [Соссюр, 1999, с. 26–27]. В силу первичности речи устанавливается следующая взаимозависимость между языком и речью: «язык одновременно и орудие, и продукт речи» [там же: 26].

Согласно Ф. де Соссюру, речь — это «сумма всего того, что говорят люди», включающая в себя «комбинации, в которых говорящий использует код (code) языка с целью выражения своей мысли» и «психофизический механизм, позволяющий ему объективировать эти комбинации» [там же, с. 21–22]. Оба эти аспекта зависят от воли говорящего, а значит, речь — в отличие от языка — не может быть коллективной. Она проявляется одномоментно, являя собой «сумму частных случаев» [там же, с. 26], и индивидуально, создавая некое отражение говорящего. Таким образом, лингвистика делится на *лингвистику языка* (которая затем стала структурной) и *лингвистику речи* (которая теперь является скорее функциональным направлением).

Ф. де Соссюр также вводит термин *речевая деятельность*, представляющий собой процесс говорения и его восприятия участниками коммуникации. Язык и речь, в свою очередь, являются сферами функционирования речевой деятельности.

В отечественном языкознании важный вклад в проблему разграничения языка и речи внёс Л.В. Щерба. Учёный выделял три следующих *аспекта языковых явлений*: процессы говорения и понимания как *речевую деятельность*, словарь и грамматику как *языковую систему* и тексты («совокупность всего говоримого и понимаемого») как *языковой материал* [Щерба, 1974, с. 26–28]. Все три аспекта, безусловно, тесно связаны между собой. Так, речевая деятельность является индивидуальным проявлением языковой системы, которая, в свою очередь, выводится из языкового материала. В то же время, речевая деятельность — это языковой материал, несущий в себе потенциальные перемены в языковой системе [там же, с. 28–29]. Таким образом, термин *речевая деятельность* несколько неоднозначен и остаётся таковым до сих пор.

А.И. Смирницкий соглашается с разделением языка и речи, предложенным Ф. де Соссюром, называя речь «сырым материалом исследования», а язык — «предметом изучения, извлекаемым из этого материала» [Смирницкий, 1957, с. 14]. Учёный говорит о том, что «язык неизбежно изучается в речи» [там же].

Л. Ельмслев, в свою очередь, выделяет четыре *аспекта языковых явлений*: *схему*, *норму*, *узус* и *речевой акт* (РА). В этом четырёхчленном делении два компонента — норма и речевой акт — имеют отношение к устной речи. Так, *норма* — материальная форма схемы в данной социальной реальности, а *речевой акт* — речь как индивидуальное пользование узусом [Ельмслев, 1960].

Немецкий ученый Э. Косериу вводит трихотомию *система — норма — речь*. При этом к устной речи он относит как речь, в которой язык «функционирует и конкретно дан» [Косериу, 1963, с. 156], так и норму — «систему обязательных реализаций» [там же, с. 198].

Согласно концепции психолингвистов Л.С. Выготского, Н.И. Жинкина и А.А. Леонтьева речевая деятельность — один из видов человеческой деятельности (наравне с трудовой, познавательной и др.), и в реальности речь представлена лишь в виде неё [Выготский, 1956]. А.А. Леонтьев отмечает, что речевая деятельность есть не только совокупность актов речи, но и связная система, которая подчиняется законам функционирования языка. Согласно учёному, в речевой деятельности каждого конкретного индивида осуществляются выделенные Р.О. Якобсоном [Якобсон, 1975] функции языка [Леонтьев, 1969]. О том, что речь не является простым отражеием системы языка, говорит и Н.И. Жинкин [Жинкин, 1958]. Таким образом, проблемы определения речевой деятельности, поднятые Л.В. Щербой, постепенно находят своё разрешение.

## Особенности разговорной речи

Большинство определений устной и письменной речи затрагивают их основное отличие — канал коммуникации. Так, устная речь — «речь звучащая, произносимая» [Трошева, 2006а, с. 567], а письменная — речь «изображённая <…> с помощью специальных графических знаков» [Трошева, 2006б, с. 285]. Таким образом, исследователь, занимающийся устной речью, имеет дело со звучащей, произносимой речью.

Для устной разговорной речи характерны такие признаки, как спонтанность, линейность развёртывания во времени, автоматизм в отборе языковых средств. Кроме того, её отличает от письменной присутствие интонации, жестов и мимики, влияющих на информативную и перформативную функции речи [Трошева, 2006б].

Восприятие разговорной речи характеризуется двусторонностью: «Отдельный акт речи, речевой акт в нормальных случаях представляет собой двусторонний процесс, охватывающий говорение и протекающие параллельно и одновременно слуховое восприятие и понимание услышанного» [Маслов, 1987, с. 11].

Значительное влияние на разговорную речь оказывает и характер адресата, в качестве которого выступает конкретный непосредственный собеседник (или собеседники) [Шмид, 2003, с. 39–108].

Л.В. Щерба отмечает, что письменный литературный и устный языки чрезвычайно далеки, однако ещё более далеки литературный и разговорный языки [Щерба, 1957а, с. 118].

На протяжении долгих лет исследователи языка занимались, в первую очередь, языком литературным — наддиалектной формой существования языка. Изучение же устной разговорной речи было на периферии в силу как труднодоступности необходимого оборудования, но также и её «более низкого социального престижа» по сравнению с литературной формой. Однако изучение разговорной формы необходимо для пополнения знаний о языке в целом: «Наблюдать мы должны жизнь, а не бессмысленно повторять то, что о ней написано в книжках» [Бодуэн де Куртенэ, 1960, с. 243]. Более того, именно устная речь является первичной формой существования языка [Гаспаров, 1978].

Л.В. Щерба также писал о необходимости исследования разговорной речи: «написанный язык живет не самостоятельно, а питается соками живого, произносимого языка — лучше же наблюдать источник жизни, нежели ее наросты; к тому же в развитии произносимого языка участвуют силы всего народа, а письменного — лишь ничтожной его части» [Щерба, 1957б, с. 14–15].

Тех же взглядов придерживался и М.В. Панов. Он писал: «звуковая речь “главнее” письменной. Каждый человек говорит больше, чем пишет (и обычно слушает больше, чем читает)» [Панов, 1979, с. 4].

Так, с течением времени становится всё очевиднее научная необходимость «повернуться “лицом к живой речи” и выработать <…> навыки ее наблюдения, записи и лингвистического анализа» [Полевая лингвистическая практика …, 2010, с. 13].

В связи с вышеизложенным в настоящей работе возникает необходимость точно определить термин *разговорная речь* (РР) и отделить его от понятия *устной речи*.

Н.Ю. Шведова свидетельствует: «далеко не всё устное, произносимое (и даже воплощающееся в разговоре) относится к речи разговорной» [Шведова, 1960, с. 3]. В то же время, О.Б. Сиротинина включает любую речь, проявляющуюсю в устной форме, в понятие *устной речи* [Сиротинина, 1983, с. 5].

В настоящей работе за *разговорную речь* принимается разновидность «устной литературной речи, обслуживающая повседневное обиходно-бытовое общение и выполняющая функции общения и воздействия» [Лингвистический эницклопедический словарь, 1990, с. 407]. Так, отличие РР от устной речи заключается в более ограниченной сфере употребления: РР используется в обиходном общении.

РР стала изучаться относительно недавно, во второй половине XX века. Вопросы выделения релевантных, определяющих характеристик РР до сих пор актуальны. Д.Н. Шмелёв писал: «Те области речи, которые обычно определяются как “устная”, “разговорная” или “устно-разговорная” речь, не изучены еще настолько, чтобы можно было бы со всей определенностью утверждать, что то или иное выделение изучаемого объекта является во всех отношениях наиболее целесообразным» [Шмелев, 1977, с. 22]. О.Б. Сиротинина отмечает неофициальность РР: «непринужденность типична лишь для неофициального общения, поэтому разговорная речь — это устная неофициальная речь» [Сиротинина, 1983, с. 20]. Е.А. Земская же противопоставляет непринуждённую устную речь другим формам — бытовой речи, просторечию, диалектной речи [Земская, 1979, с. 3]. Е.А. Земская утверждает, что важную роль в определении РР играют её основные особенности: «непринужденность отношений между партнерами коммуникации, неподготовленность речевого акта и отсутствие установки на сообщение, имеющее официальный характер» [Земская и др., 1981, с. 57]. В свою очередь, О.А. Лаптева в качестве важнейшего признака устно-разговорной речи называет непубличность всех её жанров [Лаптева, 1974а, c. 1]. Б.М. Гаспаров говорит о необратимости и многоканальном характере передачи и обработки сообщения как о важнейших свойствах устной речи [Гаспаров, 1978].

Проблемным вопросом современной лингвистики остаётся положение РР в системе русского языка. В частности, ведутся научные дискуссии о существовании не только письменной, но и устной (разговорной) нормы. Одни учёные относят РР к *кодифицированному литературному языку* (КЛЯ), другие — противопоставляют ему. Например, Е.А. Земская рассматривает РР и КЛЯ как некоторую диглоссию внутри *литературного языка* (ЛЯ) [Земская и др., 1981, с. 21–22]. Данная точка зрения принята и в настоящей работе. Б.М. Гаспаров, О.А. Лаптева и О.Б. Сиротинина относят РР к ЛЯ в качестве его разновидности, или особого стиля. О.А. Лаптева вводит понятие *устно-разговорной разновидности литературного языка* [Лаптева, 1974б; 1974в; 1975].

Итак, устная разговорная речь есть звучащая речь, используемая для непосредственного общения. В основе создания грамматики языка всегда лежат факты речи, а значит, РР (как и устная речь в целом), занимающая несравнимо большую часть речевого поведения, должна быть объектом многоаспектного научного исследования.

## Корпусные методы в исследованиях устной речи

*Корпусная лингвистика* занимается «разработкой общих принципов построения и использования лингвистических корпусов (корпусов текстов) с использованием компьютерных технологий» [Захаров, 2005, с. 3]. Эта наука появилась в конце прошлого века в связи с распространением и востребованностью методов автоматической обработки данных. Предметом корпусной лингвистики выступают «теоретические основы и практические механизмы создания и использования представительных массивов языковых данных, предназначенных для лингвистических исследований в интересах широкого круга пользователей» [там же, с. 4].

Корпусный материал может быть получен с помощью методов полевой лингвистики, которая изучает живую речь в неотрывной связи с носителем языка, учитывая при этом его социальные и психологические характеристики [Богданова, 2007]. А.Е. Кибрик определяет её как «лингвистическую дисциплину, разрабатывающую и практикующую методы получения информации о неизвестном исследователю языке на основании работы с его носителями» [Кибрик, 2017]. Учёный обращает внимание на следующее противопоставление: «Полевая лингвистика неявным образом противопоставлена “кабинетной”, для которой источником данных являются либо языковая интуиция самого исследователя <…>, либо обширный корпус текстов на изучаемом языке, о котором опять же известно достаточно много для того, чтобы изучать его без обращения к суждениям его носителей. <…> Полевая лингвистика неразрывно связана с человеком — носителем языка, который является посредником между исследователем и языком» [там же]. Методы полевой лингвистики активно используются при разработке лингвистических корпусов.

*Лингвистический*, или *языковой*, *корпус текстов* — «большой, представленный в электронном виде, унифицированный, структурированный, размеченный, филологически компетентный массив языковых данных, предназначенный для решения конкретных лингвистических задач» [Захаров, 2005, с. 3]. Перечисленные характеристики определяют основные признаки, необходимые для построения корпуса на базе массива электронных текстов. Существует два типа корпусной разметки — лингвистическая (морфемная, морфологическая, синтаксическая, семантическая и т. п.) и мета-разметка (сведения о типе текста в корпусе и его характеристиках). Таким образом, корпусная лингвистика тесно взаимодействет с *компьютерной лингвистикой* в реализациях своих основных целей.

Корпусы могут делиться на различные типы по ряду классифицирующих признаков: тип данных (письменные, речевые, смешанные корпусы), «параллельность» (одноязычные, двуязычные, многоязычные корпусы), положение в системе национального языка (литературные, диалектные, разговорные, терминологические, смешанные корпусы) и др. [Захаров, 2005, с. 13].

### Обзор наиболее известных корпусов для русского языка

В настоящее время корпусы существуют для многих национальных языков.

Наиболее известными корпусами русского языка считаются

* *Национальный корпус русского языка*[[1]](#footnote-1) (НКРЯ) (более 360 млн. токенов; включает как письменные тексты различных жанров, так и устные),
* *Упсальский корпус русского языка* (1 млн. словоупотреблений; включает публицистические и художественные произведения второй половины XX века) и созданные на его основе *Тюбингенские корпуса русских текстов*[[2]](#footnote-2) (более 26 млн. словоупотреблений; содержат художественные и публицистические тексты),
* *Хельсинкский аннотированный корпус русских текстов* ХАНКО[[3]](#footnote-3) (100 тыс. словоупотреблений из журнала «Итоги»),
* *Компьютерный корпус текстов русских газет конца ХХ века*[[4]](#footnote-4) (200 тыс. словоупотреблений),
* *Корпус русского литературного языка*[[5]](#footnote-5) (более 1 млн. словоупотреблений; представлены художественные, публицистические и научные тексты).

Существуют также и крупные корпуса, включающие интернет-тексты:

* *Генеральный интернет-корпус русского языка*[[6]](#footnote-6), созданный при помощи технологий автоматического сбора и разметки текстов новостей, блогов, комментариев в социальных сетях и т. п. (более 20 млрд. токенов),
* Русский корпус в составе многоязычного *ruTenTen corpus*[[7]](#footnote-7) (около 16 млрд. словоупотреблений),
* *Открытый корпус русского языка* (около 2 млн. токенов; включает тексты различных жанров (газетные статьи, статьи Википедии, блоги и проч.); пополняется и размечается пользователями Интернета[[8]](#footnote-8),
* *Корпус русских учебных (академических) текстов*[[9]](#footnote-9) (включает курсовые работы, выпускные квалификационные работы, эссе и т. п.; помимо традиционных видов разметки содержит разметку по ошибкам).

Кроме того, для русского языка созданы и такие корпусы, объектом фиксации которых является устная речь.

* *Корпус устных текстов* (КУТ) в *Национальном корпусе русского языка*, который, однако, составляет лишь 2.8 % от всего объема текстов НКРЯ. КУТ, в свою очередь, включает записи публичной речи, транскрипты кинофильмов, а также записи разговорной речи, сделанные Е. А. Земской, М. А. Китайгородской и Н. Н. Розановой, транскрипты речи, предоставленные А. С. Гердом и М. В. Русаковой, О. Б. Сиротининой, и расшифровки текстов, собранных научным коллективом под руководством А. С. Асиновского.
* Собрание *Рассказы о сновидениях и другие корпуса звучащей речи*[[10]](#footnote-10) — корпус, представляющий исключительно устную речь (включает четыре модуля: рассказы о сновидениях — около 14 тыс. словоупотреблений, рассказы сибиряков о жизни — около 5 тыс. токенов, веселые истории из жизни — около 10 тыс. словоупотреблений, истории о подарках и катании на лыжах — около 4,5 тыс. токенов на русском, около 4,5 тыс. словоупотреблений на армянском и около 4 тыс. токенов на японском языках. Уникальность данного корпуса состоит в наличии дискурсивной транскрипции в трёх вариантах (минимальной, упрощённой и полной), а также в выделении в качестве минимальной синтактико-дискурсивной составляющей текста элементарной дискурсивной единицы (ЭДЕ).

Наконец, на филологическом факультете Санкт-Петербургского государственного университета создается корпус, фиксирующий русскую устную речь в её диалогической и монологической разновидностях, –*Звуковой корпус русского языка* (ЗКРЯ). Данный корпус является крупнейшей базой данных по современной русской устной разговорной речи. Помимо текстов, ЗКРЯ содержит социологические характеристики информантов (пол, возраст, место рождения и проживания, уровень образования, профессиональная занятость и др.), а для некоторых информантов — и психологические характеристики, полученные с помощью анкетирования (психотип, уровень нейротизма и пр.) Целью данного проекта является исследование закономерностей разворачивания естественной устной речи. ЗКРЯ представляет собой один из первых подходов к исследованию звучащей речи с применением новых методологических возможностей и с формированием новых терминологических решений [Богданова и др, 2009, с. 38; Звуковой корпус …, 2013]. В корпусе присутствуют как монологи, записанные по разным лингвистическим программам, так и диалоги, протекающие в разных экстралингвистических условиях. Такое многообразие материала позволяет изучать различные типы и жанры УР и их характеристики в сравнении друг с другом.

В настоящей работе Звуковой корпус русского языка, включающий в себя устные тексты и их расшифровки, взят в качестве источника материала. Таким образом, данное исследование выполнено в рамках *корпусного подхода*.

### Звуковой корпус русского языка

В настоящий момент ЗКРЯ состоит из двух основных модулей — корпуса «Один речевой день» (ОРД) и *Сбалансированной аннотированной текстотеки* (корпуса, сбалансированного по различным параметрам, материал для которого собирается по *принципу ковчега* — «каждой твари по паре»[Звуковой корпус …, 2013, с. 67].

Материалом для данного исследования послужил корпус ОРД. Главная цель проекта — изучение речевого поведения носителя языка в течение дня. При этом используется метод 24-часовой записи (эта методика применяется вслед за Британским национальным корпусом, использующим тот же принцип записи устной части) [Степанова, 2008; Asinovsky et al., 2009; Богданова-Бегларян и др., 2015; 2017б]. Прототипом корпуса стал корпус японской спонтанной речи JST/CREST ESP Project, призванный решить задачу фиксации устной речи носителей языка и обучения ей многофункциональных роботизированных систем с интерфейсом «человек-машина» [Сибата, 1983].

На начальном этапе материал собирался по *принципу невода* [Богданова, 2010б]: «забрасывается широкая сеть в среду носителей языка, вытягивается всё, что в нее попало, и становится объектом многоуровневого исследования» [Звуковой корпус …, 2013, с. 68]. Позже на первый план при сборе материала выступила *социолингвистическая* балансировка, предполагающая учёт таких социальных характеристик информантов, как пол, возраст, профессиональная занятость, уровень речевой компетенции [Богданова, 2010а, с. 36] и др.

Социолингвистическая информация о говорящих доступна из анкет, в которых отражены их основные характеристики: пол, возраст, образование, профессия, место рождения и др. [Bogdanova-Beglarian et al., 2016]. На данный момент в состав ОРД входит речь информантов следующих социальных групп:

* по гендеру: мужчины и женщины;
* по возрасту: молодёжная, средняя и старшая группы;
* по профессиональной принадлежности: рабочие, инженеры, военнослужащие, представители естественных наук, представители гуманитарных наук, работники образования, представители сферы обслуживания, IT-специалисты, офисные служащие, творческая интеллигенция;
* 5 статусу: студенты и учащиеся; наемные работники и специалисты; руководящие работники; бизнесмены и частные предприниматели; неработающие, в том числе пенсионеры.

Одновременно с проведением звукозаписей осуществляется психолингвистическое тестирование информантов — выявление таких психологических характеристик, как уровень экстраверсии/интроверсии и уровень нейротизма, определяемых с помощью личностного опросника Г. Айзенка [Звуковой корпус …, 2013, с. 67].

Такой подход к подбору материала позволяет выявлять корреляции между различными лингвистическими параметрами речи и индивидуальными характеристиками говорящего, а также даёт возможность сравнивать результаты, полученные от информантов из разных социальных групп.

Объём корпуса к настоящему моменту превышает 1 млн. словоупотреблений в расшифровках и содержит 1250 часов звучания, более 2800 коммуникативных макроэпизодов, 130 информантов (из них 69 мужчин и 61 женщина в возрасте от 18 до 83 лет) и более 1000 основных коммуникантов [Богданова-Бегларян и др., 2017а].

Материал представлен в расшифрованном орфографическом виде. В транскриптах отмечаются физические паузы, а также проявления спонтанности — паузы хезитации (в том числе заполненные), обрывы, самокоррекция, — и другие паралингвистические элементы — смех, вздох, кашель и пр.

На базе речевого корпуса ОРД проводятся разнообразные исследования русского языка повседневного общения и устного русского дискурса, в том числе в социолингвистическом аспекте [Богданова-Бегларян и др., 2016; Bogdanova-Beglarian et al., 2015].

## Социолингвистический аспект в исследованиях устной речи (и синтаксиса разговорного языка)

*Социолингвистика* берёт истоки у изучения диалектов (как вариантов языка, обусловленных территориально) и арго (как языков, используемых в специфических закрытых группах). Ещё И.А. Бодуэн де Куртенэ указывал на необходимость учитывания социального аспекта в лингвистическом исследовании: «Так как язык возможен только в человеческом обществе, то <…> мы должны отмечать в нем всегда сторону социальную. Основанием языковедения должна служить не только индивидуальная психология, но и социология» [Бодуэн де Куртенэ, 1963, с. 348]. Окончательно *социолингвистика* оформилась как самостоятельная наука лишь в XX веке.

Социолингвистика — область языкознания, «изучающая язык в связи с социальными условиями его существования. Под социальными условиями имеется в виду комплекс внешних обстоятельств, в которых реально функционирует и развивается язык: общество людей, использующих данный язык, социальная структура этого общества, различия между носителями языка в возрасте, социальном статусе, уровне культуры и образования, месте проживания, а также различия в их речевом поведении в зависимости от ситуации общения» [Социолингвистика …, 2017]. Для обработки данных в социолингвистике «используются разновидности корреляционного анализа. Для социолингвистики типичны корреляции, в которых в качестве независимых переменных выступают те или иные социальные параметры, стратификационные или ситуативные, а в качестве зависимых — языковые явления. Между коррелятами отмечается как полная, так и неполная функциональная зависимость» [Лингвистический энциклопедический словарь, 1990, с. 481]. По этому принципу выделяются два типа экстралингвистических факторов, влияющих на речь: влияние переменных компонентов *коммуникативной ситуации* и относительно постоянного *социального статуса* говорящих.

*Коммуникативная ситуация* — ситуация речевого общения, включающая двух или более участников. Д.Х. Хаймс описал структуру коммуникативной ситуации на основе следующих социолингвистических параметров:

1. S=SETTING/SCENE (обстановка) — время и место, в которых существует речевой акт, а также физические обстоятельства в целом;
2. P=PARTICIPANTS (участники) — говорящий(-е) и слушающий(-е);
3. E=ENDS (цели) — задачи и цели, а также результаты РА;
4. A=ACT SEQUENCE (последовательность речевых актов) — форма речевых событий и их порядок в цепи РА;
5. K=KEY (указания на эмоциональную сферу) — тон и манера протекания РА;
6. I=INSTRUMENTALITIES (код и канал) — коды и каналы передачи сообщения;
7. N=NORMS (нормы) — социальные конвенции, использующиеся собеседниками для достижения коммуникативной цели;
8. G=GENRE (речевой жанр) — жанры речевого общения [Hymes, 1974, с. 55].

По мнению У. Брайта, самые важные факторы, влияющие на языковые различия — социальная принадлежность Отправителя и Получателя сообщения, а также обстановка, в которой имеет место РА. Притом социальная принадлежность говорящего «ярче всего проявляется в случае “классовых диалектов”, когда различия в речи соответствуют различиям в социальной стратификации. <…> Сюда же относятся и случаи различия между “мужской” и “женской” речью» [Брайт, 1975, с. 35].

*Социальный статус* — «соотносительная позиция индивида или группы, определяемая социальными признаками (экономическое положение, профессия, квалификация, образование и т. п.), природными признаками (пол, возраст и т. д.), а также престижем и местом в структуре власти» [Социология …, 2003, с. 789]. В понятие социального статуса разные исследователи включают различные характеристики, как собственно социальные, так и биологические. В.И. Карасик, отмечая междисциплинарный характер понятия, даёт ему следующую характеристику: «1) общественно значимые различия между людьми имеют биологическую и социальную природу и типизируются в системе социальных признаков индивида, 2) социальные признаки индивида образуют иерархию в соответствии с ценностями определённого сообщества в определённый период, 3) социальные признаки индивида неоднородны в различных отношениях, группируются в признаковые комплексы и могут быть измерены, 4) социальный статус человека раскрывается в ролевых, дистанционных и нормативных характеристиках поведения, обусловленных социально-экономическими и культурно-этическими факторами общественной жизни, 5) ролевые, дистанционные и нормативные характеристики поведения отражаются в стратегиях и средствах невербального и вербального общения» [Карасик, 1991, с. 322].

Социолингвистические характеристики говорящего определённым образом влияют на лингвистические особенности его речи. Отсюда вытекает необходимость изучения *социолектов*. Социолект — «совокупность языковых особенностей, присущих какой-либо социальной группе — профессиональной, сословной, возрастной и т. п. — в пределах той или иной подсистемы национального языка» [Беликов, Крысин, 2001, с. 33]. Группу людей, использующих социолект, объединяет характерный репертуар выбираемых ими языковых средств. Тем не менее, речь любой группы, говорящей на определённом социолекте, можно отнести к любому коду в системе данного национального языка. Границы же групп в отношении социолектов могут и, как кажется, должны быть размытыми и множественными, так как в норме каждый человек владеет несколькими социолектами. Следует отметить, что термин *социолект* имеет в некоторой степени двойственный характер: с одной стороны, это «социальный тип, который проявляется у человека под влиянием черт, свойственных данной расе, этнической группе, национальности, социальному классу» [Ерофеева Т.И., 2002, с. 174], с другой — «система речевых средств определенной группы, детерминированая рядом факторов, имеющих не только социальный, но и биологический и психологический характер (например, пол, возраст, темперамент и т. д.)» [там же].

Социолиектами, а также языковыми группами, выделяемыми по социальному статусу их членов, занимается *макросоциолингвистика*. Отмечается, что изучаемые группы — «это группы условные» [Беликов, Крысин, 2001, с. 216], члены таких групп могут не иметь контактов между собой. Макросоциолингвистика выявляет и учитывает наборы социально значимых признаков, по которым образуется социолект.

У. Лабов отмечает, что социолингвистика описывает язык с учётом социальных характеристик говорящего [Labov, 2006, с. 7]. О том, какие именно характеристики оказывают наибольшее влияние на речь, существуют различные мнения. М.В. Панов выделяет в качестве ведущих факторов возраст, место рождения, место постоянного жительства, профессия, социальное положение родителей, место рождения родителей [Панов, 1990]. А.С. Штерн считает основными чертами пол и возраст как биологические основы личности [Штерн, 2006, с. 131]. Л.П. Крысин отмечает три аспекта, которые влияют на распределение языковых вариантов среди носителей национального языка: возрастной, профессиональный и социально-территориальный [Крысин, 1989]. Т.И. Ерофеева и Е.В. Ерофеева отмечают, что с особенностями речи коррелируют пол, возраст, место рождения, уровень образования и специальность [Ерофеева Т.И., 2004; Ерофеева Е.В., 2005]. Н.В. Богданова в числе социальных параметров, определяющих репертуар и функционирование единиц в речи, выделяет пол, возраст, профессиональную принадлежность, профессиональное или непрофессиональное отношение к речи и уровень речевой компетенции [Богданова и др., 2008].

В настоящей работе предпринимается попытка определить, влияют ли социальные характеристики говорящего на синтаксический уровень языка (в частности, на дистантные связи предикатных групп) и социальными характеристиками говорящего. В данном исследовании в ходе социолингвистического анализа материала учитываются следующие факторы: пол, возраст и профессиональная принадлежность информантов.

### Пол говорящего как социолингвистический фактор

Пол — постоянная социобиологическая характеристика человека. В некоторых исследованиях используется также термин *гендер* — «социокультурный конструкт, связанный с приписыванием индивиду определённых качеств и норм поведения на основе его биологического пола» [Кирилина, 2005, с. 112].

Гендер говорящего может проявляться в речи на любом языковом уровне. В первую очередь, для речи мужчин и женщин характерно употребление различных лексических единиц. Хотя в эпоху феминизма некоторые отличия были в значительной степени стёрты, тем не менее, существует вероятность выбора определённого языкового варианта человеком, принадлежащим к определённому гендеру [Labov, 1971]. Кроме того, этот выбор во многом определяется культурными стереотипами языкового сообщества, к которому относит себя говорящий.

Д. Таннен, изучая коммуникативные неудачи в общении коммуникантов разного пола, описывает мужской и женский *гендерлекты*. Различия связаны с коммуникативными стратегиями и тактиками, различных у мужчин и женщин. Поведение мужчин, в целом, нацелено на достижение результата и сохранение независимости и социального статуса. Поведение женщин, в свою очередь, нацелено в большей степени на выстраивание и сохранение хороших отношений с собеседником и выражение эмоций [Tannen, 1990].

П.М. Фишман выявляет неравный вклад мужчин и женщин в построение коммуникации. Женщины чаще задают вопросы, используют фатические сигналы, вводящие тему клаузы. Учёный приходит к выводу о том, что именно женщины выстраивают разговор так, чтобы минимизировать коммуникативные неудачи [Fishman, 1978, с. 404–405].

П. Традгилл и Е.Б. Райен отмечали, что мужчины более склонны к использованию нестандартного языка, а женщины — к использованию престижного варианта [Ryan, 1979, с. 155; Trudgill, 1983]. Н.Б. Вахтин и Е.В. Головко также пишут об этой особенности: «…обычно все инновации попадают в язык через мужскую речь» [Вахтин, Головко, 2004, с. 76].

Е.А. Земская, М.А. Китайгородская и Н.Н. Розанова провели исследование, посвящённое особенностям мужской и женской речи (в семейной повседневной обстановке). Было обнаружено, что для женщин характерно включать в разговор бытовую тематику, что вызвано экстралингвистическими факторами, например, обстановкой и действиями говорящих. По мнению исследователей, это скорее связано не с полом, а с социальными и семейными функциями женщины. На речь мужчин в большей степени влияет фактор профессиональной деятельности (для них характерно употребление в речи специальных терминов). Кроме того, в мужской речи чаще встречаются эскпрессивные, стилистически сниженные средства и бранная лексика (которую женщины также используют, однако только в разговоре с другими женщинами). Для женской речи характерна экспрессивность другого типа: использование слов-интенсивов, гиперболизации, выражающих удивление междометий.

Различаются, безусловно, и семантические поля. Типично мужскими оказываются спорт, охота, профессиональная сфера и военная сфера; типично женскими же — природа, животные, эмоциональная сфера межличностного взаимодействия и окружающий быт. Тем не менее не стоит забывать о том, что границы между мужской и женской речью могут быть проницаемыми, подвижными и размытыми; а выделенные особенности могут быть тенденциями, связанными не с полом, а с иными факторами [Земская и др., 1993].

Т.В. Гомон разделяет признаки мужской и женской речи на поверхностные и глубинные. К поверхностным признакам относятся тематические. К глубинным признакам мужской речи — использование армейского и тюремного жаргона, газетно-публицистических клише, вводных слов со значением констатации, а также нецензурных слов как вводных, употребление большого количества абстрактных существительных, преобладание нецензурных инвектив и конструкций, обозначающих действия и процессы, гглаголов активного залога и переходных, однообразие лексических приёмов при передаче эмоций и пр. К глубинным признакам женской речи относятся: множество вводных слов, определений, обстоятельств, местоименных подлежащих и дополнений, а также модальных конструкций, выражающих различную степень неуверенности, неопределённости, склонность к употреблению «престижных» форм, клише, книжной лексики, эффемизмов, употребление оценочных высказываний с дейктическими лексемами вместо имени, использование конструкций *наречие + наречие*, двойного отрицания, высокая эмоциональная окраска [Гомон, 1990].

Гендерные различия в речи социальных групп проявляются неоднородно. Е.И. Горошко в ходе психолингвистических экспериментов показала, что в наименьшей степени они обнаруживаются у лиц с высшим образованием, занимающихся интеллектуальной деятельностью [Горошко, 1999].

Т.И. Ерофеева на материале Пермского корпуса живой речи доказывает, что гендерный аспект коррелирует с некоторыми фонетическими, лексическими и синтаксическими чертами устных текстов. Такими чертами являются супрасегментные признаки речи (логическое ударение, эмоциональное ударение, темп речи), употребление архаизмов и жаргонизмов, «длина предложения» [Ерофеева Т.И., 2002, с. 174].

На материале модуля ОРД ЗКРЯ также были выявлены различия в речи мужчин и женщин. Женщины в целом говорят больше (противоположные данные даёт С. Зимин [Zimin, 1981]), но медленнее мужчин (схожие данные предлагает С.Б. Степанова [Степанова, 2011, с. 44]), а также используют меньше хезитации и вопросительных предложений, но больше прагматем [Богданова-Бегларян, 2014; 2016] и редуплицированных форм [Попова, 2016]. В женской речи чаще встречается редуцированная форма *щас*, а в мужской — *чё*.

Исследование частоты встречаемости лексических единиц в повседневной русской речи в зависимости от условий коммуникации было проведено Т.Ю. Шерстиновой [Шерстинова, 2016]. Женщины больше внимания уделяют формальному поддержанию разговора (*угу, хорошо*), мужчины же больше хезитируют, а в верхней зоне их частотного словаря появляются бранные слова. Самым частотным словом в мужской речи оказалась частитца *ну*, в женской — личное местоимение *я* [там же, с. 622–623].

### Возраст говорящего как социолингвистический фактор

В социолингвистике возраст, как и пол, понимается не только как биологическая характеристика. Пол говорящего — это сложное явление, которое количественно определяется постоянными биологическими факторами и качественно — социальными (в связи с принадлежностью к какому-либо поколению) и психологическими (которые в некоторой степени влияют на личность в каждый определённый период) [Ананьев, 1968, с. 226; Аванесов, 1984, с. 9; Грановская, 1988].

Возрастные группы могут быть разделены различным способом. Например, психолог Э. Эриксон описывает восемь фаз психосоциального развития человека: младенчество, раннее детство, игровой (дошкольный) и школьный возраст, юность, молодость, взрослость и старость [Эриксон, 2000]. Академия педагогических наук в 1965 году приняла более дробное деление: новорождённый, грудной ребёнок, раннее детство, затем два периода детства, подростковый и юношеский возраст, два периода среднего возраста, пожилой человек, старческий возраст и долгожитель [Венгер, 1971, с. 270]. А.С. Штерн, в свою очередь, после собственно детства выделяет раннюю и позднюю юность, раннюю, среднюю и позднюю зрелость и оценку прошлой жизни [Штерн, 2006, с. 131–132].

В исследованиях речи пристальное внимание обращено на средний возраст. Е.В. Ерофеева, опираясь на анализ устной речи, выделяет три периода внутри среднего возраста: раннюю (от 25 до 34 лет), среднюю (от 25 до 44 лет) и позднюю (от 45 до 55 лет) зрелость [Ерофеева Е.В., 2005, с. 131].

Онтолингвистика — область возрастной лингвистики (науки о проявлении возрастных характеристик в языке говорящего и изменении языка с возрастом) — изучает становление речи человека в детстве. Данным вопросом занимались такие учёные, как С.Л. Рубинштейн [Рубинштейн, 1940], Л.С. Выготский [Выготский, 1956], Н.И. Жинкин [Жинкин, 1958], Т.А. Ладыженская [Ладыженская, 1986; 1875], А.Р. Лурия [Лурия, 1979], С.Н. Цейтлин [Цейтлин, 2000], К.Ф. Седов [Седов, 2004] и др.

В отличие от особенностей детской речи, черты речи говорящих старшего и среднего возраста описаны в научной литературе в значительно меньшей степени. Тем не менее, можно говорить о том, что данный аспект тесно связан с понятием нормы в языке. Зачастую именно различное понимание нормы представителями разных возрастных групп становится причиной «конфликта поколений».

Речь говорящих старшего возраста более консервативна, насыщена выходящими из употребления элементами. Молодёжь же привносит в язык инновации, экспрессию. Язык людей средней возрастной группы отличается наибольшей нормативностью: «В нормально развивающемся языке, опирающемся на язык трех поколений, “стандарт” определяется с ориентацией на среднее поколение: именно оно является носителем того варианта языка, который в данном обществе принимается за норму» [Вахтин, Головко, 2004, с. 78].

Кроме того, «статистически значимыми оказались и различия в темпе речи информантов младшей и старшей возрастных групп (до и после 40 лет), с возрастом мы говорим медленнее: 6.1 vs.5.3 сл./с» [Степанова, 2011, с. 45].

В качестве релевантных выступают и акустические признаки голоса, позволяющие определить речь старшего поколения по особому произношению и дополнительным шумам, возникающим из-за возрастных изменений. Существуют различные гипотезы на этот счёт. Одна из них гласит, что разрушение речевых норм в пожилом возрасте само по себе является нормой [Coupland N., Coupland J., Giles, 1991, с. 1–24]. Другая объясняет возрастные изменения возвращением человека «во второе детство» [Якобсон, 1996]. Более того, в первую очередь утрачиваются синтаксические структуры — то, что дети усваивают в последнюю очередь.

Противоположной точки зрения придерживается О.В. Ильичева, говоря о том, что язык говорящих из старшей возрастной группы сложнее и разнообразнее, чем у людей с высоким УРК [Ильичева, 1998, с. 33].

Т.И. Ерофеева в своей монографии отмечает, что возраст оказывает наибольшее влияние на лексический уровень речи [Ерофеева Т.И., 1991], а также на среднюю длину «предложения» и сложность синтаксиса (в старшей возрастно группе они становятся меньше) [Ерофеева Т.И., 2012, с. 55]. На других языковых уровнях возраст признаётся несущественным фактором.

Е.В. Ерофеева, анализируя фонетическую сторону речи пермяков, отмечает влияние возраста на частоты таких показателей, как ёканье, качественная редукция /u/, утрата и ослабление смычки у аффрикаты /с/ и других [Ерофеева Е.В., 2009, с. 23].

На материале корпуса «Один речевой день» также проводились исследования лингвистических параметров, на которые влияет возраст говорящего. Так, речь молодёжи характеризуется наиболее высоким темпом речи (5.91 сл/с), количеством вопросительных реплик (2.5 vs. 1.3 % у старшей группы), более частым использованием эмотивных междометий, неологизмов, жаргонизмов и бранной лексики, а также использованием непроективных (инверсированных) конструкций. В речи говорящих среднего возраста меньше всего синтаксических нарушений, что подтверждает наибольшую стандартизированность речи людей среднего возраста. В речи информантов старшего возраста зачастую присутствуют прагматемы, а также обрыв фразы [Русский язык повседневного общения …, 2016, с. 201–205].

В настоящей работе принимается возрастная периодизация, включающая в себя младшую (до 30 лет), среднюю (от 31 года до 45 лет) и старшую возрастные группы (от 46 лет).

### Профессия говорящего как социолингвистический фактор

М.М. Бахтин отмечал, что «все слова пахнут профессией, жанром, <…> определённым человеком, поколением, возрастом, днем и часом» [Бахтин, 1975, с. 106]. Без сомнения, профессинальная занятость оказывает большое влияние на язык, который использует человек. Очевидно, что «люди интеллектуального труда <…> как правило, используют литературную форму национального языка, крестьяне пользуются диалектом, а те, кто занят физическим трудом, наряду с литературным языком могут прибегать к средствам просторечия, профессиональных жаргонов» [Беликов, Крысин, 2001, с. 162].

Л.П. Крысин определяет *профессиональные языки* как «языковые особенности, которые связаны со спецификой профессий» [Крысин, 1977, с. 29]. Учёный делит данные особенности на терминологические и жаргонные: «каждая профессия имеет своеобразные средства выражения, известные людям только этой профессии и не употребляющиеся за ее пределами, причем это не только слова-обозначения специальных вещей и действий (т. е. термины), но и разнообразные способы речи, выработанные людьми общей профессии и служащие им в качестве своего языка, а точнее — своего профессионального жаргона» [там же, с. 33].

Наряду с профессиональным жаргоном представители социальных групп используют и другую форму языка — разговорную речь. РР также неизбежно подвергается влиянию профессионального жаргона, шире — профессиональной деятельности. В то время как «за пределами соответствующей профессиональной среды профессиональные жаргонизмы непонятны» [Словарь социолингвистических терминов, 2006, с. 176], они всё же встречаются в спонтанной речи людей в общении не только внутри профессионального коллектива, но и вне его. Н.С. Арапова определяет *профессионализмы* (профессиональные жаргонизмы) как «слова и выражения, свойственные речи представителей той или иной профессии или сферы деятельности, проникающие в общелитературное употребление (преимущественно в устную речь) и обычно выступающие как просторечные, эмоционально окрашенные эквиваленты терминов» [Лингвистический энциклопедический словарь, 1990, с. 403]. Так, профессионализмы могут со временем потерять узкий профессиональный характер и стать общепонятными, перейдя в категорию общеупотребительной лексики.

Научный коллектив под руководством Н.В. Богдановой-Бегларян исследовал влияние профессии на бытовую речь медиков и юристов и выявил, что «медиков “выдает” пристрастие к медицинской терминологии, они прибегают к ней тем чаще, чем выше уровень речевой компетенции (УРК) информанта» [Богданова, 2010а, с. 36], а юристов — «проникновение элементов официально-делового стиля, столь характерного для юридического языка, в бытовые рассказы об отдыхе» [там же, с. 37].

Перцептивные эксперименты показали, что носители русского языка весьма успешно соотносят предложенные отрывки речи с определёнными профессиональными группами, в частности, юристов [Иванова, 2008; 2010]. В речи же информантов-философов О.А. Иванова обнаружила обилие терминов, связанных с психологической жизнью человека [Иванова, 2011].

Т.И. Ерофеева показала значимость специальности как фактора, влияющего на речь на лексическом уровне. По её мнению, профессия влияет на знание, употребление и реализацию диалектных, архаических и жаргонных лексем [Ерофеева Т.И., 2002, с. 174]. Е.Б. Русакова отметила, что социолекту «компьютерщиков» свойственна особая пространственно-временная организация, склонность к языковому творчеству, взаимодействие с англиским языком, экспрессивный характер и др. [Русакова, 2007], а А.И. Малиновская указала, что эти особенности нередко проявляются в их РР [Малиновская, 2012].

Кроме того, некоторые профессиональные группы склонны к употреблению в бытовой речи определённых типов синтаксических конструкций. Так, в речи частных предпринимателей преобладают незавершённые структуры, а в речи руководителей они не встречаются, в речи учащихся чаще всего встречаются эллиптические и парцеллированные структуры, неработающие пенсионеры почти не используют самокоррекцию [Богданова-Бегларян и др., 2017а].

# ГЛАВА 2. СИНТАКСИЧЕСКИЕ ОСОБЕННОСТИ РУССКОЙ РАЗГОВОРНОЙ РЕЧИ

## Синтаксис устной и письменной речи

Как было отмечено выше, устная и письменная речь имеют ряд существенных отличий. Каждая из форм имеет свои особенности на всех языковых уровнях. Так, письменная форма отличается от устной более сложным синтаксисом, большим количеством отвлечённой и терминологической лексики. Особенно заметны эти особенности на синтаксическом уровне.

Синтаксису письменной речи свойственны сложные и осложнённые предложения. Как следствие, большое значение в ней имеет порядок слов, а также логическая последовательность в изложении мыслей. С другой стороны, для письменной речи характерны сложные синтаксические конструкции, причастные и деепричастные обороты, распространённые определения, вставные конструкции и т. д.

Синтаксис же устной речи, в свою очередь, упрощён за счёт преобладания простых, коротких «предложений». Однако он в то же время более непредсказуем за счёт отсутствия таких жёстких правил, как в письменной речи. Порядок слов в спонтанной речи может быть совершенно свободным, иногда даже, как кажется, нелогичным. Однако эти его особенности вытекают из требований устной коммуникации [Гаспаров, 1978]. Несомненно, разговорный синтаксис нуждается в самостоятельном изучении и описании.

## Синтаксические исследования русской разговорной речи

Одними из первых в советской лингвистике разговорный синтаксис начали описывать Е.А. Земская [Земская, 1979], О.А. Лаптева [Лаптева, 1974а], О.Б. Сиротинина [Сиротинина, 1983].

О.Б. Сиротинина определяет разговорную речь как «устную форму диалогической речи бытового или нейтрального типа разговорного стиля» [Сиротинина, 1969, с. 386]

О.А. Лаптева пишет о таких признаках разговорной речи, как диалогичность, спонтанность, непринуждённость, неофициальность, непосредственность, эмоциональность, ситуативность, тематическая обусловленность, системность, устная форма, ограниченность обиходно-бытовой сферой и др. [Лаптева, 1974а, с. 6]. Данные признаки обусловлены тем, что устно-разгворная разновидность русского языка, по О.А. Лаптевой, может и должна реализоваться только в условиях коммуникативного задания повседневной речевой практики [Лаптева, 1974а, с. 11]. Речь системно организована: ей свойственен достаточно ограниченный набор специфических языковых единиц — в том числе и синтаксических конструкций — «которые обычно выступают комплексно, обусловливая друг друга в своём появлении в речи, и выстраиваются во взаимосвязанные упорядоченные ряды» [Лаптева, 1974а, с. 16]. То, что ситуативность, использование жестов и мимики, особая роль интонации и линейный характер ведут к формированию собственного набора единиц со специфическим их функционированием утверждает и Е.А. Земская [Земская, 1979, с. 69–71].

На синтаксическом уровне для устно-разговорной речи характерны «построения случайного характера, незавершённая организация, возникающие при самоперебивах, прервании высказывания, оговорках, повторениях, перестройках речи и т. д.» [Лаптева, 1974а, с. 34]. Отклонения от кодифицированного синтаксиса в устной речи не только допускаются, но в некоторых случаях и нормативно узакониваются. Например, нормальны для разговорной речи конструкции бессоюзного подчинения, именительного в позиции косвенного, «двойные глаголы» (напр., *лежу читаю, сидят ждут*).

Специфические черты разговорного синтаксиса возникают в связи с такими принципами организации устного высказывания, как принцип отсутствия предварительного обдумывания, принцип чередования ударных и безударных звеньев, принцип превалирования семантического плана высказывания над формальным. Вследствие влияния всех этих факторов в спонтанной речи возникают клишированные построения и типизированные конструкции. Они становятся синтаксической моделью, представляющей собой скопление структурно-функциональных признаков. Впоследствии такая модель может обрести многочисленные реализации и модификации.

Одно из существующих описаний системы аспектов предложения принадлежит Л. Теньеру [Теньер, 1988]. Учёный выделяет четыре аспекта: структурный, линейный, категориальный и функциональный. По Л. Теньеру, весь структурный синтаксис зиждется на отношениях, существующих между структурным и линейным порядками [там же, с. 19].

Синтаксические аспекты изучают такие разделы синтаксиса, как категориальный синтаксис (синтаксис частей речи), функциональный синтаксис (синтаксис членов предложения), сочетательный синтаксис (синтаксис связей), линейный синтаксис (синтаксис порядка слов) и интонационный синтаксис. Это поле исследований можно назвать микросинтаксисом [Богданов, 1977, с. 147].

Наряду с микросинтаксисом существует и макросинтаксис. Макросинтаксис — это синтаксис связного текста. Его основными объектами являются межфразовые единства и структура сверхфразовых единств.

В.Г. Адмони предполагает, что смысловая самодостаточность высказывания — т. е. его оторванность и от речевой ситуации, и от линейного контекста — рождается не столько за счёт лексики, сколько за счёт грамматики, прежде всего — синтаксиса. Учёный делит высказывания на «разовые», характеризующиеся «усечённостью» состава при обязательном наличии «сочетательной проекции» по отношению к предыдущему высказыванию, и генеретивные, функционирующие как цельные, самодостаточные и воспроизводимые [Адмони, 1994].

С теоретической точки зрения О.А. Лаптева призывает «обратиться к установлению реального состава современного русского литературного языка в его чрезвычано сложном и многообразном функционировании как общенационального языка с разветвлённой системо разновидностей, обеспечивающей ему возможность осуществлять задачи общения в современном языковом коллективе с его весьма дифференцированными социальными, производственными, бытовыми, культурными и иными градациями» [Лаптева, 2013, с. 6].

С практической точки зрения синтаксический анализ является неотъемлемым этапом автоматической обработки как текста, так и речи. В результате такой обработки «устанавливается, удовлетворяет ли цепочка словоформ, образующих предложение, структурным условиям, заданным явным образом правилами грамматики языка» [Буторов, 1996, с. 142].

Синтаксис также представляет интерес при исследовании речи с позиций социолингвистики. Идиостиль говорящего может проявляться на всех уровнях языка. Так, фонетические, морфологические и лексические особенности речи могут отражать диалектные особенности, связанные с прошлым или настоящим местожительством говорящего. Морфологический уровень также весьма важен в психолингвистическом аспекте. С социолингвистической же точки зрения наиболее перспективным является анализ синтаксиса, стилистики, дискурса. Г.Я. Мартыненко пишет, что «в синтаксическом строе наиболее откровенно проявляется авторская манера развёртывания мысли» [Мартыненко, 1996, с. 436].

В исследованиях П.М. Эйсмонт рассматривается синтаксис спонтанного нарратива, который наряду с семантикой является средством выражения цельности и связности текста [Эйсмонт, 2008]. В частности, было проведено сравнение структур и средств связи высказываний в речи детей и взрослых [Эйсмонт, 2007; Eysmont, 2010].

1. 1.
	2.

Изучение синтаксиса устной речи на базе корпуса «Один речевой день» проводилось такими учёными, как Г.Я. Мартыненко, П.В. Реброва и Т.Ю. Шерстинова. Г.Я. Мартыненко выдвинул гипотезу о левосторонности устной речи (подробнее о ней см. п. 2.4. данной работы) [Bogdanova et al., 2015; Martynenko, Yadchenko, 2015; Мартыненко, 2015; Мартыненко и др., 2016; Русский язык повседневного общения …, 2016, с. 120–129]. Исследования П.В. Ребровой [Реброва, 2014; 2017] выполнены с опорой на работы учёного. Т.Ю. Шерстинова же исследует синтаксис устной речи через призму N-грамм анализа, в частности, были проанализированы позиции начала-конца фраз и синтагм [Шерстинова, 2018].

## Подходы к анализу синтаксических структур

З.К. Тарланов говорит о том, что современное языкознание в России дрейфует в сторону концептологии, выдвинувшейся на передний край научного интереса и поиска [Тарланов, 2005, с. 172–181].

Европейские синтаксические теории многочисленны и различны: от сугубо логических (где главенствует предложение, рассматриваемое как форма формы, т. е. двойная форма) до формальных (где систематизируются лексико-морфологические формы, образующие предложение). Результатом стал слишком большой терминологический аппарат. Различные направления и школы вынуждены создавать работы не только о языке, но и о своём метаязыке.

В последнее время замечается всеобщий интерес к концептам как к феноменам, организующим и конституирующим глубинное семантическое пространство языка. При этом сам термин «концепт» весьма неоднозначен, так как разными исследователями понимается по-разному. Несомненно, проследить концепт проще всего на лексико-семантическом уровне. Однако методика вычленения и систематизации концептов как постоянных элементов предметно-понятийной сферы может быть перенесена и на синтаксис. Так, Г.А. Волохина и З.Д. Попова полагали, что «синтаксический концепт существует на уровне общечеловеческих смыслов и часто служит опорой для переводчиков» [Волохина, Попова, 2003, с. 8]. На лексико-семантическом уровне языка прослеживаются этнокультурные концепты, в то время как на синтаксическом уровне — надэтнические, межкультурные, общечеловеческие. В таком понимании знаком синтаксического концепта можно считать структурную схему. Тем не менее, «синтаксис важен и интересен не столько как концептовыражающая сфера, сколько как сфера, в известной мере определяющая модус восприятия мира, окружающего человека», а «синтаксический тип предложения в его парадигматической соотнесённости содержателен сам по себе, как материальный знак» [Тарланов, 2005, с. 179]. Наконец, З.К. Тарланов полагает, что вводить термин концепт в синтаксический анализ можно только при условии, что под него подводятся только типы предложения как таковые [там же].

Также в лингвистике популярна теория дискурса. Возникновение и развитие этой теории связано с желанием вывести синтаксис за пределы предложения (например, гиперсинтаксис Палека [Palek, 1968]), с разработкой прагматического аспекта речи (речевые акты) и с подходом к речи как к социальному действию (понятие перформатива). Под текстом стала пониматься формальная, абстрактная конструкция, а под дискурсом — различные виды её актуализации.

Как кажется, на уровне синтаксиса формальный подход можно признать оптимальным. По крайней мере, ни одна из эвристических — интуитивных — процедур не стала инструментом исследования значительных массивов языкового материала, за исключением работы М.Н. Петерсона [Петерсон, 1923].

## Симметрия дистантных связей в устной и письменной речи

Г.Я. Мартыненко была выдвинута гипотеза о принципиальном отличии синтаксиса устной речи от письменной с точки зрения симметрии линейных порядков [Мартыненко, 2015].

В ряде исследований было показано, что в письменной речи зависимые члены предложения имеют тенденцию к симметричному расположению слева и справа от глагола-сказуемого (в частности, [Шепелёв, 2009]).

На материале подвыборки корпуса «Один речевой день» Г.Я. Мартыненко и его ученики показали, что в русской устной речи наблюдается тенденция к существенному преобладанию левых связей над правыми, т. е. глагол тяготеет к концу высказывания (что находит своё подтверждение в работах Ю.О. Салматовой [Салматова, 2010] и П.В. Ребровой [Реброва, 2014]).

Таким образом, можно утверждать, что письменные жанры (поэзия, художественная и научная проза) стремятся к зеркальной симметрии, а устная речь, в свою очередь, тяготеет к золотой симметрии, т. е. к симметрии на основе золотого сечения [Мартыненко, 2015, с. 377]. В устной речи процент левоветвящихся глагольных групп возрастает до 59%, количество симметричных и одноэлементных (состоящих только из глагола) групп составляет по 14%, при этом процент правоветвящихся глагольных групп падает до 13%. Таким образом, по Г.Я. Мартыненко, русская устная речь является левосторонней [Bogdanova et al., 2015, с. 433–434]. Проверка этой гипотезы составляет одну из задач настоящего исследования.

# ГЛАВА 3. ИССЛЕДОВАНИЕ ДИСТАНТНЫХ СВЯЗЕЙ В ПРЕДИКАТНЫХ ГРУППАХ РУССКОЙ РАЗГОВОРНОЙ РЕЧИ

1.

## Материал и методика

Общим материалом настоящего исследования послужил речевой корпус «Один речевой день» (ОРД) — крупнейший ресурс, предназначенный для изучения разговорной речи, разрабатываемый в Санкт-Петербургском государственном университете и содержащий более 1250 часов звучания. Корпус содержит более 2800 коммуникативных макроэпизодов, записанных от 130 информантов и более 1000 их основных коммуникантов. Объём текстовых расшифровок ОРД составляет более 1 миллиона словоупотреблений [Богданова-Бегларян и др., 2017, с. 77–78]. Материал корпуса может быть разделён по 20-ти социальным группам (2 гендерных — мужчины и женщины; 3 возрастных — молодёжная (младшая), средняя, старшая; 10 профессиональных — рабочие, инженеры, военнослужащие, представители естественных наук, представители гуманитарных наук, работники образования, представители сферы обслуживания, IT-специалисты, офисные служащие, творческая интеллигенция), а также по 5-ти статусным группам (студенты и учащиеся; наемные работники и специалисты; руководящие работники; бизнесмены и частные предприниматели; неработающие, в том числе пенсионеры). Объём лингвистически проаннотированного на синтаксическом уровне материала ОРД — более 12 тысяч структур [там же].

Для настоящей работы из корпуса «Один речевой день» была извлечена подвыборка из 11 коммуникативных эпизодов, отражающих регулярные ситуации речевого общения и содержащие речь 10 информантов-представителей разных социальных групп. Материал был сбалансирован по полу и возрасту, кроме того была предпринята попытка по возможности максимально представить в выборке разные профессиональные группы. В данной подвыборке было выделено и проаннотировано 830 предикатных групп, которые и стали предметом данного исследования.

Состав информантов с учётом их социальных характеристик:

1. по полу (о влиянии гендерной принадлежности на речь см. раздел 1.4.1. настоящей работы):
	* 5 мужчин;
	* 5 женщин;
2. по возрасту (о влиянии возраста на речь см. раздел 1.4.2. данного исследования):
	* I группа (до 30 лет) — 3 информанта (20, 24 и 28 лет);
	* II группа (от 31 года до 45 лет) — 3 информанта (31, 35 и 45 лет);
	* III группа (от 46 лет) — 4 информанта (48, 53, 62 и 67 лет);
3. по профессиональной принадлежности (о влиянии профессиональной занятости говорящего на его речь см. раздел 1.4.3. настоящего исследования):
	* 3 работника сферы образования;
	* 2 IT-специалиста;
	* 2 офисных работника;
	* 1 представитель гуманитарных наук;
	* 1 представитель силовых структур;
	* 1 представитель творческих профессий.

Анализ данного материала был проведён с помощью традиционных лингвистических методов (описательного, сопоставительного и др.), при формировании количественных подсчётов использовались статистические методы анализа данных.

На *первом этапе* работы были отобраны макроэпизоды для дальнейшего исследования с учётом социальных характеристик информантов.

На *втором этапе* исследования был составлен рабочий подкорпус исследовательского материала. Из отобранных макроэпизодов были выделены предикатные группы (ПГ), впоследствии проаннотированные вручную, а также автоматически проаннотированные с помощью морфологического парсера TreeTagger. Для каждой ПГ в рабочем подкорпусе содержится информация о социальных характеристиках её автора-информанта.

На *третьем этапе* работы с помощью квантитативного статистического методов были составлены частотные таблицы количественных характеристик ПГ и их ранговое распределение — по полу, возрасту и профессиональной принадлежности.

На *четвёртом этапе* исследования на основе проанализированного материала были составлены частотные таблицы структур ПГ и их ранговые распределения для всех социальных групп в целом и для каждой из них в отдельности. Была осуществлена проверка основной гипотезы настоящей работы о левосторонности русской устной речи.

Таким образом, в результате многоаспектного анализа подкорпуса были выявлены некоторые количественные и качественные (частеречные) характеристики предикатных групп русской разговорной речи.

### Линейный синтаксис

Цепочки слов, используемые в спонтанной речи, носят линейный характер. Это обусловлено тем, что речь разворачивается во времени, которое, в свою очередь, не только одномерно, но и однонаправленно. Это объясняет линейное построение речи. Таким образом, исследование спонтанной УР с позиций линейного синтаксиса представляется не только возможным, но и первоочередным.

Для такого изучения речи необходимо выделить в ней такие синтаксические единицы, которые могут быть выделены линейно и подвержены анализу. Линейный аспект отражает «линейную организацию предложения, то есть порядок слов» [Богданов, 1977, с. 146], а «в результате взаимодействия препозиции и постпозиции с цепочками последовательно зависимых слов возникают линейные последовательности с левым и правым последовательным подчинением» [Мартыненко, 1996, с. 441]. Анализ таких последовательностей представлен, в частности, в работах П.В. Ребровой [Реброва, 2014].

### Понятия предикатной группы и синтаксического элемента

Синтаксической единицой письменной речи является предложение. Однако оно выделяется графически, что делает невозможным применение термина «предложение» к устной речи. Для обозначения синтаксических единиц УР учёные используют различные понятия: высказывания [Бахтин, 1986, с. 258], клаузы (напр., [Тестелец, 2001, с. 229–268]), элементарные дискурсивные единицы (ЭДЕ) [Кибрик, Подлесская, 2009, с. 55–61] и др.

В настоящем исследовании с этой целью используется термин *предикатная группа* (ПГ), под которым понимается предикат и окружающие его синтаксические элементы. В отличие от традиционного для лингвистики термина *предикативная группа*, определяемого как предикат и зависимые от него единицы, в ПГ в данной работе включаются также формально не зависимые от него элементы (частицы, междометья и пр.), тем не менее относящиеся к цепочке слов ПГ. Например,

*ну посчитайте там господи /* (женщина, 43 года, офисный работник).

Под синтаксическим элементом понимается графическое слово, а также такие случаи сращения графических слов, которые нельзя разделить вставкой ещё одного слова. Например,

*я имею в виду ...* (мужчина, 24 года, офисный работник).

За один синтаксический элемент в настоящем исследовании приняты следующие сращения:

* + - как бы

*ну как бы это ничего не говорит //* (женщина, 20 лет, представитель гуманитарных наук);

* + - иметь в виду

*имей в виду //* (мужчина, 53 года, работник сферы образования);

* + - на самом деле

*мало чем отличится / на самом деле / от обычного чемпионата //* (мужчина, 24 года, офисный работник);

* + - в смысле

*в смысле а чего он х... () хочет сделать то ?* (женщина, 20 лет, представитель гуманитарных наук);

* + - вряд ли

*но вряд ли ему это будет интересно /* (мужчина, 53 года, работник сферы образования);

* + - потому что

*потому что стиральная машина то грязная //* (женщина, 48 лет, IT-специалист).

Заминки, обрывы и неречевые хезитативы не учитывались в работе, так как имеют лишь косвенное отношение к синтаксической структуре высказывания. Паузы были учтены при делении ПГ в сложных случаях. Например,

*бабочки //* ***\*П*** *тоже очень сезонно //* (женщина, 28 лет, работник сферы образования).

### Синтаксические категории

Для аннотирования синтаксических элементов за основу были взяты правила кодирования высказываний, предложенные П.В. Ребровой [Реброва, 2014, с. 66–67]. Данные правила синтаксической разметки были модифицированы таким образом, чтобы отвечать задачам настоящего исследования. Таким образом, была получена классификация категорий, для которой используются следующие обозначения:

D – дискурсивные слова;

F – фразеологизмы;

Inf – инфинитив;

N – отрицание;

Q – вопросительные слова;

S – имя существительное, местоимение;

V – глагол, спрягаемая форма;

Y – согласие;

Z – остальные частицы;

А – имя прилагательное;

В – обстоятельство;

М – обращение;

Н – отрицательная частица «не»;

О1 – прямое дополнение;

О2 – косвенное дополнение;

О3 – дополнение с предлогом.

### Методы формального подсчёта

Для анализа данных были применены квантитативные методы анализа в частности, статистические. Из отобранного материала был сформирован рабочий подкорпус предикатных групп с указанием говорящих и их социальных характеристик.

Затем каждая ПГ была проаннотирована вручную, после чего был подсчитан объём (количество синтаксических элементов, составляющих ПГ), левая и правая дистантность каждой группы (т. е. количество элементов до и после предиката соответственно). Таким образом, были выявлены количественные характеристики устных «предложений», что дало возможность их анализа. С помощью статистических методов были выявлены наиболее частотные по количественным характеристикам конструкции в речи всех информантов в целом и каждой социальной группы, представленной в подкорпусе, в частности.

Анализ структуры предикатных групп проводился с помощью программы — морфологического парсера TreeTagger. Все ПГ были автоматически проаннотированы, после чего была проведена ручная корректировка результатов. Полученные данные были преобразованы в частотные таблицы. Таким образом, было создано 12 ранжированных списков частеречных структур ПГ: 1) для всех говорящих; 2) для информантов-мужчин; 3) для информантов-женщин; 4) для информантов до 30 лет; 5) для информантов от 31 года до 45 лет; 6) для информантов от 46 лет; 7) для представителейгуманитарных наук; 8) для IT-специалистов; 9) для работников сферы образования; 10) для офисных работников; 11) для представителей силовых структур и 12) для представителей творческих профессий.

На основе полученных данных была подтверждена гипотеза о левосторонности устной речи, а также сделан ряд выводов относительно возможных тенденций в речи представителей той или иной социальной группы. Данные выводы были проверены с помощью статистических методов (критерий Стьюдента, критерий Фишера), что позволило выявить недостоверные с математической точки зрения заключения и определить те, которые можно признать статистически значимыми.

## Результаты

1.
2.
3. 1.
	2.

Количественные характеристики дистантных связей предикатных групп русской устной речи, выявленные в результате исследования, были сопоставлены с аналогичными данными для письменной речи. Таким образом, был проведён сравнительный анализ дистантных связей, характерных для устного и письменного каналов коммуникации. Кроме того, данные о дистантностях ПГ были собраны для каждой социальной группы с целью выявления характерных различий.

### Дистантные связи в устной и письменной речи

С помощью количественного и статистического анализа было определено, что средний объём ПГ равен 4.28 элемента. Левая дистантность в среднем равна 2.49 элемента, правая — 0.8 элемента (см. таблицу 1). Так, гипотеза о левосторонности русской устной речи, выдвинутая Г.Я. Мартыненко, подтверждается на материале настоящего исследования.

*Таблица 1*

*Описательные статистики предикатных групп устной спонтанной речи*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 830 | 4.280723 | 1.000000 | 17.00000 | 2.230827 |
| Левая дистантность | 830 | 2.487952 | 0.000000 | 11.00000 | 1.855250 |
| Правая дистантность | 830 | 0.796386 | 0.000000 | 10.00000 | 1.353452 |

Таким образом, среднестатистическая ПГ спонтанной устной речи в количественном отношении может быть представлена так:

***2:х:1***,

где

2 — количество элементов до предиката,

х — предикат,

1 — количество элементов после предиката.

Подобные таблицы были составлены не только для всех говорящих, но и для каждой социальной группы в отдельности (см. таблицы 9-19 в приложении). Полученные данные позволяют вычислить типичные в количественном отношении ПГ для каждой социальной группы, представленные в таблице 2.

*Таблица 2*

*Схемы типичных ПГ для каждой социальной группы*

|  |  |
| --- | --- |
| Социальная группа | Схема ПГ |
| Мужчины | *2:х:1* |
| Женщины | *3:х:1* |
| Молодёжная возрастная группа (до 30 лет) | *3:х:1* |
| Средняя возрастная группа (31-45 лет) | *2:х:1* |
| Старшая возрастная группа (от 46 лет) | *3:х:1* |
| Гуманитарии | *3:х:1* |
| Работники сферы образования | *3:х:1* |
| IT-специалисты | *2:х:1* |
| Офисные работники | *2:х:1* |
| Представители силовых структур | *3:х:0* |
| Представители творческих профессий | *2:х:1* |

Очевидно, что левосторонние структуры типичны для речи всех социальных групп. Они составляют 70.12% (582 ПГ) от исследуемого материала, в то время как симметричные и правосторонние структуры — лишь 15.42% (128 ПГ) и 14.46% (120 ПГ) соответственно. Статистическая значимость разницы в количестве левоветвящихся структур по сравнению с другими была проверена с помощью критерия Фишера: эмпирическое значение φ оказалось равно 24.51 и 23.98 для пар «левоветвящиеся» vs. «правоветвящиеся» и «левоветвящиеся» vs. «симметричные» соответственно при критическом значении 2.31 (при p ≤ 0.01). Таким образом, полученные данные являются статистически значимыми и подтверждают гипотезу о левосторонности русской устной речи.

### Дистантные связи в речи различных соцгрупп

Полученные результаты были также сопоставлены внутри каждой социальной категории: по гендеру, по возрасту и по профессиональной сфере информантов.

Исследование показало, что средний объем ПГ у мужчин равен 4.29 элемента, левая дистантность составляет 2.44 элемента, правая — 0.85 элемента. Аналогичные данные для женской речи следующие: 4.28 элемента, 2.52 элемента и 0.76 элемента соответственно. Так, значимая зависимость количественных характеристик ПГ от пола говорящего не прослеживается (см. рисунок 1).

*Рисунок 1*

*Количественные характеристики ПГ в различных гендерных группах*

При анализе отличий в размере ПГ для говорящих из разных возрастных групп (см. рисунок 2) было выявлено, что в средней возрастной группе средний объем ПГ оказался самым маленьким (3.9 элемента), а в молодёжной группе — самым большим (4.67 элемента).

*Рисунок 2*

*Количественные характеристики ПГ в различных возрастных группах*

При этом на гистограмме видно, что разница в объёме ПГ у всех возрастных групп создаётся исключительно за счёт левой дистантности при почти одинаковой правой (0.8, 0.81 и 0.76 элемента для младшей, старшей и средней групп соответственно).

Для определения достоверности данного вывода был использован статистический метод расчёта критерия Стьюдента. При сопоставлении выборок, представляющих молодёжную и среднюю возрастные группы, t-критерий Стьюдента оказался равен 4. Данное эмпирическое значение входит в зону значимости при её критическом уровне p ≤ 0.01 (см. рисунок 3), а значит, выводы, сделанные на основе сравнения двух данных возрастных групп, являются достоверными.

*Рисунок 3*

*Проверка значимости выводов, сделанных на основе соотношения младшей и средней возрастных групп, при критическом уровне значимости p ≤ 0.01*

Наибольшие отличия были выявлены в речи представителей различных профессиональных групп (см. рисунок 4). Максимальный средний объём ПГ ожидаемо оказался у работников образования (4.8 элемента) и гуманитариев (4.7 элемента) по сравнению с информантами других профессий (от 3.69 элемента до 4.03).

*Рисунок 4*

*Количественные характеристики ПГ в различных профессиональных группах*

Хотя объём материала не позволяет с уверенностью говорить о каких-либо закономерностях, о которых могут свидетельствовать выявленные отличия, однако уже на данном этапе можно предположить, что профессиональная принадлежность оказывает определённое влияние на количественные характеристики единиц речи говорящего.

Ранговое распределение частотных по объёму ПГ также позволило увидеть определённые закономерности.

Самая частотная по объему предикатная группа у информантов-мужчин состоит из 5 элементов, вторая по частоте — из 4 элементов, третья — из 3. В то время как у информантов-женщин картина обратная (см. рисунок 5).

*Рисунок 5*

*Ранговое распределение частотных по объёму ПГ в различных гендерных группах*

Однако судить о том, тенденция это или случайность, следует, опираясь на больший объём материала.

Сравнение частотных по объёму ПГ в речи информантов различных возрастных групп показало весьма интересный результат. Самая частотная ПГ в молодёжной возрастной группе состоит из 5 элементов, в то время как самые частотные ПГ и в средней, и в старшей возрастных группах состоят из 3 элементов (см. рисунок 6).

*Рисунок 6*

*Ранговое распределение частотных по объёму ПГ в различных возрастных группах*

Ранговое распределение объёма ПГ по частотности вообще сходно у информантов от 31 года (представителей средней и старшей возрастных групп), что позволяет выделить молодёжь как тяготяющую к более широким предикатным группам. Примеры показывают, что их размер достигается чаще всего за счёт дискурсивных слов:

*то есть как бы даже если я поднимаю руку* (женщина, 20 лет, гуманитарий)

Однако следует также проверить данный вывод с помощью критерия Стьюдента. Достоверность выводов, сделанных на основе сравнения младшей и средней возрастных групп, была доказана выше (см. рисунок 3).

При сопоставлении же молодёжной и старшей возрастных групп эмпирическое значение t-критерия Стьюдента оказалось вне зоны значимости (см. рисунок 7). Таким образом, сделанный вывод необходимо скорректировать, что и было сделано (см. п. 3.3. настоящего исследования).

*Рисунок 7*

*Проверка значимости выводов, сделанных на основе соотношения младшей и старшей возрастных групп*

Что же касается сопоставления средней и старшей возрастных групп, то эмпирическое значение t-критерия Стьюдента в данном случае находится в зоне неопределённости (см. рисунок 8). То есть сделанный при их сравнении вывод, возможно, достоверен, однако считать его окончательным представляется преждевременным. Это означает, что сходство между средней и старшей возрастными группами, также не доказано.

*Рисунок 8*

*Проверка значимости выводов, сделанных на основе соотношения средней и старшей возрастных групп*

Наибольшие различия в размере ПГ были выявлены в речи разных профессиональных групп говорящих (см. рисунок 9).

*Рисунок 9*

*Ранговое распределение частотных по объёму ПГ в различных профессиональных группах*

Полученные результаты, тем не менее, следует проверить на большей выборке материала, так как при вычислении их достоверности по критерию Стьюдента эмпирические значения оказываются в зоне значимости лишь при сопоставлении тех профессиональных групп, которые представлены наибольшим количеством материала. Профессиональные группы показали самые значительные различия, однако делать выводы об их конкретных особенностях представляется преждевременным, так как материала по некоторым группам недостаточно для получения достоверных данных.

## Выводы

В связи с вышеизложенным следующие выводы, сделанные на основе количественного анализа, можно считать достоверными.

1. Речь всех исследуемых социальных групп является левосторонней.
2. Объём ПГ меняется в первую очередь за счёт левой дистантности, в то время как правая дистантность у всех информантов вне зависимости от пола, возраста и профессий колеблется от 0 до 1 элемента.
3. Количественные характеристики ПГ в русской разговорной речи, по всей видимости, не зависят от пола говорящего.
4. Средний объём ПГ в речи представителей молодёжной возрастной группы больше, чем в речи представителей средней.

Для того, чтобы доказать или опровергнуть остальные выводы, сделанные в п. 3.2, следует провести исследование на бо́льших объёмах речевого материала.

Однако два первых вывода подтверждают основную гипотезу работы о преобладании левых дистантных связей над правыми. Таким образом, настоящее исследование подтверждает положение о том, что для синтаксических единиц (в частности, предикатных групп) устной речи характерна левосторонность в отличие от зеркальной симметрии в письменной речи.

# ГЛАВА 4. POS-АНАЛИЗ ПРЕДИКАТНЫХ ГРУПП

1.

## Частота встречаемости различных типов ПГ

Особый интерес представляют собой не только количественные, но и качественные характеристики предикатных групп. Прежде всего, как кажется, стоит обратиться к структуре ПГ, в частности, частеречной.

Исследуемый материал был автоматически проаннотирован с помощью программы — морфологического парсера TreeTagger, после чего была выполнена ручная проверка и корректировка результатов аннотирования.

Так, был получен набор из 830 частеречных структур, каждая из которых принадлежит информанту, обладающему известным набором социальных характеристик.

Одним из очевидных способов классификации представляется деление предикатных групп на 2 типа:

1. глагольные ПГ;
2. неглагольные ПГ.

За глагольные в данном исследовании принимаются такие ПГ, которые имеют глагол (в любой форме) в качестве ядра. Например,

*там ещё один остался //* (мужчина, 24 года, офисный работник).

Под неглагольными понимаются ПГ, имеющие другие части речи (например, слова категории состояния, именная часть составного именного сказуемого) в качестве ядра. Например,

*там нет моих тапочек под тобой ?* (женщина, 48 лет, IT-специалист).

Среди анализируемых предикатных групп оказалось 658 глагольных ПГ (79%) и 172 неглагольных ПГ (21%), то есть их соотношение для всех говорящих — 4:1.

Особый интерес представляют отхождения от этого соотношения в каких-либо социальных группах. Для того, чтобы их увидеть, была составлена таблица (см. таблицу 3).

*Таблица 3*

*Соотношение глагольных и неглагольных ПГ в различных социальных* *группах*

|  |  |  |  |
| --- | --- | --- | --- |
|  | V | PRAEDIC | всего |
| по возрасту | количество ПГ | доля в % | количество ПГ | доля в % | количество ПГ |
| молодёжная группа |
| 205 | 78.54406 | 56 | 21.45594 | 261 |
| средняя группа |
| 271 | 79.51807 | 68 | 20.48193 | 339 |
| старшая группа |
| 180 | 79.29515 | 48 | 20.70485 | 228 |
| по гендеру | мужчины |
| 247 | 74.39759 | 85 | 25.60241 | 332 |
| женщины |
| 411 | 82.37705 | 87 | 17.62295 | 498 |
| по профессии | представители гуманитарных наук |
| 81 | 82.65531 | 17 | 17.34694 | 98 |
| IT-специалисты |
| 123 | 80.39216 | 30 | 19.60784 | 153 |
| работники сферы образования |
| 187 | 76.63934 | 57 | 23.36066 | 244 |
| офисные работники |
| 177 | 89.39394 | 21 | 10.60606 | 198 |
| представители силовых структур |
| 30 | 61.22449 | 19 | 38.77551 | 49 |
| представители творческих профессий |
| 60 | 68.18182 | 28 | 31.81818 | 88 |

Из таблицы 3 видно, что количество глагольных ПГ у женщин несколько выше, чем у мужчин (82% против 74%), а эти же показатели в различных возрастных категориях оказались почти равны (81% у старшей группы, 79% у молодёжной и средней). Наибольшая разница в количестве глагольных и неглагольных ПГ обнаруживается среди представителей разных профессиональных групп (от 61% глагольных предикатных групп у представителей силовых структур до 89% у офисных работников).

Однако профессиональные группы являются самыми малочисленными в исследуемой выборке, поэтому относиться к касающимся их выводам необходимо с особой осторожностью. По этой причине было принято решение проверить достоверность сделанных выводов с помощью статистических методов. Для этого использовались такие математические расчёты как критерий Стьюдента и критерий Фишера.

При проверке вывода о разнице в распределении различных типов ПГ между гендерными социальными группами выяснилось, что он может считаться достоверным: эмпирическое значение t по критерию Стьюдента для этих групп равно 3.1 при критическом значении 2.56 (при p ≤ 0.01), а эмпирическое значение φ по критерию Фишера равно 2.78 при критическом значении 2.31 (при p ≤ 0.01), соответственно, оба эмпирических значения находятся в зоне значимости, так как они больше критических значений (см. рисунок 10).

*Рисунок 10*

*Проверка значимости выводов, сделанных на основе соотношения мужской и женской гендерных групп, при критическом уровне значимости p ≤ 0.01*

Проверка вывода о разнице в распределении различных типов ПГ между возрастными социальными группами показала, что он не может считаться достоверным. Критерии Стьюдента и Фишера позволяют сравнивать лишь две выборки, поэтому возрастные и профессиональные группы необходимо сравнивать попарно. Итак, эмпирическое значение t по критерию Стьюдента равно 0.6 и 0.3 для пар младшая-средняя и средняя-старшая соответственно при критическом значении 2.58 (при p ≤ 0.01), а эмпирическое значение φ по критерию Фишера равно 0.425 и 0.292 для тех же пар соответственно при критическом значении 2.31 (при p ≤ 0.01), а значит, все эмпирические значения находятся вне зоны значимости (см. рисунок 11).

*Рисунок 11*

*Проверка значимости выводов, сделанных на основе соотношения младшей, средней и старшей возрастных групп, при критическом уровне значимости p ≤ 0.01*

При проверке вывода о разнице в распределении различных типов ПГ между профессиональными социальными группами выяснилось, что он может считаться достоверным исключительно для пары «представители силовых структур» vs. «офисные работники» (см. рисунок 12).

*Рисунок 12*

*Проверка значимости выводов, сделанных на основе соотношения профессиональных групп представителей силовых структур и офисных работников при критическом уровне значимости p ≤ 0.01*

## Типичные структуры ПГ

Кроме того, был составлен ранжированный список из 610 обнаруженных частеречных структур (см. таблицу 20 в приложении). Доля 9 структур составила более 1% от всего материала (см. таблицу 4).

*Таблица 4*

*Частотные частеречные структуры ПГ*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № | структура | коли-чество | доля в % | ранг | схема ПГ | пример |
| 1 | *V* | 45 | 5.42 | 1 | *х* | сходите // |
| 2 | *S-PRO V* | 17 | 2.05 | 2 | *1:х* | я поняла // |
| 3 | *PART V* | 16 | 1.93 | 3 | *1:х* | не знаю |
| 4 | *CONJ S-PRO V* | 15 | 1.81 | 4 | *2:х* | что ты ноешь ? |
| 5 | *PRAEDIC* | 13 | 1.57 | 5 | *х* | прикольно // |
| 6 | *S V* | 12 | 1.45 | 6 | *1:х* | экспорт накроется // |
| 7 | *CONJ V* | 10 | 1.2 | 7 | *1:х* | если будет |
| 8 | *ADV-PRO V* | 10 | 1.2 | 7 | *1:х* | сейчас () посмотрю / |
| 9 | *PART PART V* | 9 | 1.08 | 8 | *2:х* | ну не знаю / |

 Таким образом, были выявлены самые частотные частеречные структуры ПГ. По типу среди них 1 неглагольная и 8 глагольных. По содержанию самая частотная структура ПГ состоит из одного глагола. 3й, 7й и 9й ранги имеют структуры, состоящие из глагола со служебными частями речи. 2й и 4й ранги принадлежат структурам, в которых к глаголу добавляется субъект действия.

Оказалось также, что объём самых частотных структур составляет всего 1-2 элемента, а их правая дистантность неизменно равна нулю. Таким образом, подтверждается основная гипотеза данного исследования о левосторонности русской устной речи.

В рамках структур ПГ также представляется интересным произвести сопоставительный анализ по различным социальным группам. Это было сделано следующим образом.

Для каждой социальной группы был составлен рейтинг из 10 самых частотных частеречных структур. Эти рейтинги были сопоставлены в рамках каждой социальной категории — пола, возраста и профессии — с помощью таблиц, представленных ниже (см. таблицы 5-8).

*Таблица 5*

*Соотношение рейтингов частотных частеречных структур ПГ в различных гендерных группах*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ранг | мужчины | доля в % | женщины | доля в % |
| 1 | *V* | 5.12 | *V* | 5.62 |
| 2 | *PRAEDIC* | 2.71 | *PART V* | 3.01 |
| 3 | *S-PRO V* | 2.11 | *S-PRO V* | 2.01 |
| 4 | *V S* | 1.81 | *CONJ S-PRO V* | 2.01 |
| 5 | *CONJ S-PRO V* | 1.51 | *S V* | 1.61 |
| 6 | *ADV-PRO V* | 1.51 | *PART PART V* | 1.61 |
| 7 | *S V* | 1.2 | *CONJ V* | 1.2 |
| 8 | *CONJ V* | 1.2 | *PART S-PRO V* | 1.2 |
| 9 | *V V* | 1.2 | *ADV-PRO V* | 1 |
| 10 | *S-PRO PART V* | 0.6 | *S-PRO PART V* | 1 |

Сравнивая гендерные группы, представляется возможным сделать следующие выводы:

* + 1. структура *V* является самой частотной во всех гендерных группах;
		2. лишь структуры *V* и *S-PRO V* имеют одинаковые (I и III соответственно) ранги в речи мужчин и женщин;
		3. в свою очередь, структуры *PRAEDIC* и *V S* в речи мужчин имеют высокие (II и IV соответственно) ранги, в то время как в речи женщин эти структуры вовсе не входят в 10 самых частотных структур (XI и XVII ранги соответственно).

Полученные данные были проверены на достоверность с помощью критерия Фишера. Было определено эмпирическое значение для соотношения каждой представленной в выводах структуры в мужской и женской речи. Оказалось, что ни один из них не может считаться достоверным, поскольку ни одно эмпирическое значение не находится в зоне значимости (см. рисунок 13).

*Рисунок 13*

*Проверка значимости выводов, сделанных на основе соотношения мужской и женской гендерных групп*

Однако третий вывод, касающийся структур *PRAEDIC* и *V S*, находится в зоне неопределённости (между критическими значениями при p ≤ 0.05 и при p ≤ 0.01), что даёт основания полагать, что он отчасти достоверен. Тем не менее, причисление этого вывода к окончательным представляется преждевременным.

*Таблица 6*

*Соотношение рангов частотных частеречных структур ПГ в различных возрастных группах*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ранг | молодёжная группа | доля в % | средняя группа | доля в % | старшая группа | доля в % |
| 1 | *V* | 4.18 | *V* | 6.78 | *V* | 4.39 |
| 2 | *S-PRO V* | 3.42 | *PRAEDIC* | 2.95 | *CONJ S-PRO V* | 2.63 |
| 3 | *PART V* | 2.28 | *PART V* | 2.36 | *S V* | 1.75 |
| 4 | *ADV-PRO V* | 1.52 | *CONJ S-PRO V* | 2.06 | *S-PRO V* | 1.75 |
| 5 | *S V* | 1.14 | *V S* | 1.77 | *CONJ V* | 1.32 |
| 6 | *PART PART V* | 1.14 | *S V* | 1.47 | *PART PART V* | 1.32 |
| 7 | *CONJ V* | 1.14 | *V V* | 1.47 | *PRAEDIC* | 0.88 |
| 8 | *PART S-PRO V* | 1.14 | *S-PRO V* | 1.18 | *PART V* | 0.88 |
| 9 | *PART S-PRO PART V* | 1.14 | *ADV-PRO V* | 1.18 | *V S* | 0.88 |
| 10 | *CONJ S-PRO V* | 0.76 | *CONJ V* | 1.18 | *ADV-PRO V* | 0.88 |

Соотношение 10 самых частотных частеречных структур ПГ в различных возрастных группах также позволяет сделать некоторые предварительные выводы:

1. структура *V* является самой частотной во всех возрастных группах;
2. структура *PRAEDIC*, имеющая II и VII ранги в средней и старшей возрастных группах соответственно, не представлена в топ-10 молодёжной группы (XVII ранг);
3. а структура *PART PART V*, имеющая VI ранг в младшей и старшей возрастных группах, не представлена в топ-10 средней группы;
4. структура *CONJ S-PRO V* тем выше в ранговом распределении, чем старше возраст говорящих (X, IV и II ранги в трёх возрастных группах соответственно).

Данные выводы также были проверены на достоверность с помощью критерия Фишера. Было определено эмпирическое значение для соотношения каждой представленной в выводах структуры в младшей и средней и в средней и старшей возрастных группах. Оказалось, что лишь вывод, касающийся структуры *PRAEDIC*, может считаться достоверным для пары «молодёжь» vs. «средняя группа» (см. рисунки 14-15).

*Рисунок 14*

*Проверка значимости выводов, сделанных на основе соотношения младшей и средней возрастных групп*

*Рисунок 15*

*Проверка значимости выводов, сделанных на основе соотношения средней и старшей возрастных групп*

Однако вывод, касающийся структуры *PRAEDIC*, для пары «средняя группа» vs. «старшая группа» находится в зоне неопределённости (между критическими значениями при p ≤ 0.05 и при p ≤ 0.01), что даёт основания полагать, что этот вывод также отчасти достоверен, хотя и не включён в данной работе в окончательные выводы (см. п. 4.3. настоящего исследования).

*Таблица 7*

*Соотношение рейтингов частотных частеречных структур ПГ в различных профессиональных группах (1)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ранг | ГУМ | доля в % | IT | доля в % | ОБР | доля в % |
| 1 | *PART V* | 6.12 | *V* | 6.54 | *CONJ S-PRO V* | 3.28 |
| 2 | *V* | 4.08 | *CONJ S-PRO V* | 2.61 | *CONJ V* | 2.05 |
| 3 | *S-PRO V* | 4.08 | *S V* | 1.96 | *V* | 1.64 |
| 4 | *PART PART V* | 2.04 | *PART V* | 1.31 | *S V* | 1.64 |
| 5 | *ADV PART V* | 2.04 | *S-PRO V* | 1.31 | *PART PART V* | 1.23 |
| 6 | *S V* | 1.02 | *PART PART V* | 1.31 | *S-PRO PART V* | 1.23 |
| 7 | *PRAEDIC* | 1.02 | *PART V S-PRO* | 1.31 | *ADV-PRO V* | 1.23 |
| 8 | *ADV-PRO V* | 1.02 | *V S* | 1.31 | *PART S-PRO V* | 1.23 |
| 9 | *PART V S-PRO* | 1.02 | *S-PRO PART V* | 1.31 | *CONJ S-PRO PART V* | 1.23 |
| 10 | *PART S-PRO V* | 1.02 | *ADV PRAEDIC* | 1.31 | *S-PRO V* | 0.82 |

*Таблица 8*

*Соотношение рейтингов частотных частеречных структур ПГ в различных профессиональных группах (2)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ранг | ОФ | доля в % | СИЛ | доля в % | ТВОР | доля в % |
| 1 | *V* | 11.62 | *ADV-PRO V* | 4.08 | *PRAEDIC* | 7.95 |
| 2 | *S-PRO V* | 3.54 | *V V* | 4.08 | *V* | 4.55 |
| 3 | *PART V* | 3.03 | *PART V* | 2.04 | *V S* | 3.41 |
| 4 | *CONJ V* | 1.52 | *CONJ S-PRO V* | 2.04 | *V V* | 2.27 |
| 5 | *CONJ S-PRO V* | 1.01 | *V S* | 2.04 | *S-PRO V* | 2.27 |
| 6 | *S V* | 1.01 | *PRAEDIC* | 2.04 | *S V* | 2.27 |
| 7 | *S-PRO PART V* | 1.01 | *PART S-PRO PART V* | 2.04 | *ADV-PRO V* | 1.14 |
| 8 | *ADV-PRO V* | 1.01 | *V ADV* | 2.04 | *V ADV* | 1.14 |
| 9 | *PART S-PRO V* | 1.01 | *S-PRO PR S V* | 2.04 | *S-PRO PR S V* | 1.14 |
| 10 | *V PR S* | 1.01 | *ADV* | 2.04 | *ADV* | 1.14 |

Соотношение 10 частотных частеречных структур ПГ в различных профессиональных группах также выявляет некоторые отличия, которые можно сформулировать как следующие предварительные выводы:

1. две самые частотные для всех структуры *V* и *S-PRO V* не встречаются у представителей силовых структур более 1 раза;
2. имеющая в общем распределении III ранг структура *PART V* не входит в топ-10 у работников сферы образования и представителей творческих профессий;
3. имеющая в общем распределении IV ранг структура *CONJ S-PRO V* не входит в топ-10 у гуманитариев и представителей творческих профессий;
4. имеющая в общем распределении V ранг структура *PRAEDIC* не входит в топ-10 у IT-специалистов, работников сферы образования и офисных работников.

Предложенные выводы также были проверены на значимость с помощью критерия Фишера. Было определено эмпирическое значение для соотношения каждой представленной в выводах структуры в каждой паре тех профессиональных групп, которых касается данное наблюдение. Оказалось, что все выводы достоверны лишь частично: первый — для структуры *V* и только при сопоставлении «силовиков» с IT-специалистами и офисными работниками, второй — только для пар «работники сферы образования» vs. «гуманитарии», «работники сферы образования» vs. «офисные работники» и «представители творческих профессий» vs. «гуманитарии», третий — только для работников сферы образования по сравнению с гуманитариями и с представителями творческих профессий, четвёртый — только для представителей творческих профессий в сопоставлении с IT-специалистами, работниками сферы образования и офисными работниками (см. рисунки 16-20).

*Рисунок 16*

*Проверка значимости выводов о структуре V*

*Рисунок 17*

*Проверка значимости выводов о структуре S-PRO V*

*Рисунок 18*

*Проверка значимости выводов о структуре CONJ S-PRO V*

*Рисунок 19*

*Проверка значимости выводов о структуре PART V*

*Рисунок 20*

*Проверка значимости выводов о структуре PRAEDIC*

Также в зоне неопределённости (между критическими значениями при p ≤ 0.05 и при p ≤ 0.01) находятся выводы о структуре *V* в сравнении представителей силовых структур и гуманитариев и представителей творческих профессий, о структуре *S-PRO V* для пары «силовики» vs. «офисные работники», о структуре *PART V* для пары «офисные работники» vs. «представители творческих профессий» и о структуре *CONJ S-PRO V* в сопоставлении IT-специалистов и гуманитариев и представителей творческих профессий. Такое положение данных выводов едва ли позволяет в настоящем исследовании считать их достоверными, поэтому в окончательные выводы (см. п. 4.3 настоящего исследования) они не включены.

И всё же такое большое количество значений в зоне значимости и неопределённости по критерию Фишера для профессиональных социальных групп позволяет говорить о наибольшей степени влияния профессиональной принадлежности на частеречную организацию ПГ в повседневной речи.

## Выводы

Анализ данных позволил сделать ряд выводов, которые были проверены с помощью статистических методов. Согласно им достоверными могут считаться следующие наблюдения.

1. Объём самых частотных структур составляет всего 1-2 элемента
2. Самые частотные структуры являются левосторонними (их правая дистантность равна 0).
3. Количество глагольных предикатных групп у женщин больше, чем у мужчин.
4. Наибольшая разница в количестве глагольных и неглагольных предикатных групп была обнаружена среди представителей разных профессиональных групп (от 61% глагольных предикатных групп у представителей силовых структур до 89% у офисных работников).
5. Структура *PRAEDIC* характерна для средней возрастной группы (II ранг) и не характерна для молодёжной (XVII ранг).
6. Структура *V* является самой частотной в профессиональных группах IT-специалистов и офисных работников, а в речи представителей силовых структур в исследуемом материале такая структура отсутствует.
7. Структура *S-PRO V* является второй по частотности в речи офисных работников, а в речи представителей силовых структур она не представлена.
8. Структура *PART V* имеет высокие ранги в профессиональных группах представителей гуманитарных наук и офисных работников (I и III соответственно), в то время как у работников сферы образования она встречается лишь 1 раз, а у представителей творческих профессий не представлена.
9. Структура *CONJ S-PRO V* является самой частотной в профессиональной группе работников сферы образования, в то же время в группах представителей гуманитарных наук и представителей творческих профессий она не представлена.
10. Структура *PRAEDIC* является самой частотной в речи представителей творческих профессий, а в речи работников сферы образования, IT-специалистов и офисных работников она встречается редко.

Гипотеза о левостороннем характере дистантных связей предикатных групп русской разговорной речи вновь оказывается подтверждена, т. к. согласно второму выводу самые частотные структуры являются левоветвящимися.

Следует отметить, что полученные на данном этапе выводы социолингвистического характера представляются несколько разрозненными; они не являются достаточными для выявления диагностических признаков исследуемых социальных групп, однако проведённое исследование демонстрирует потенциал использованных методов. Несомненно, в продолжение работы такие диагностические признаки могут быть и будут выявлены.

# ЗАКЛЮЧЕНИЕ

Настоящее исследование посвящено изучению дистантных связей в предикатных группах русской разговорной речи. Источником материала для исследования послужил Звуковой корпус русского языка повседневного общения «Один речевой день». В качестве предмета исследования в работе принята предикатная группа, определяемая как предикат и его окружение.

В первой главе работы освещены основные научные подходы к изучению устной разговорной речи. В ней приведён обзор некоторых лингвистических корпусов, а также описан социолингвистический аспект изучения «живой» устной речи.

Вторая глава исследования посвящена синтаксису устной разговорной речи, вопросам, касающимся анализа синтаксических структур и влиянию социолингвистического фактора на данный уровень языка. Представлена гипотеза о различии в симметрии дистантных связей устной и письменной речи.

В третьей главе работы отражён ход исследования. Описан способ отбора материала и методы его анализа. Приводится система синтаксических обозначений, использованных для аннотирования материала. Количественные характеристики предикатных групп устной речи статистически оцениваются для всех говорящих и для каждой социальной группы в отдельности. В работе подробно рассмотрены дистантные связи в речи различных гендерных, возрастных и профессиональных групп. На основе их сопоставления приводится ряд наблюдений. В качестве выводов представлены лишь те из них, достоверность которых была доказана с помощью строгих статистических методов. Подтверждена гипотеза о левосторонности устной речи с точки зрения дистантных связей предикатных групп.

Четвёртая глава затрагивает структурные характеристики предикатных групп. Их тип и частеречный состав анализируются так же, как и их количественные показатели в третьей главе. Сделанные наблюдения также подвергаются проверке с помощью статистических методов, в качестве выводов приведены лишь статистически значимые. Вторично подтверждена гипотеза о том, что для русской разговорной речи типичны левоветвящиеся конструкции.

Двухаспектный анализ предикатных групп позволил сделать несколько статистически значимых наблюдений, которые были отражены в качестве выводов в конце обеих практических глав. Полученные данные подтвердили левосторонний характер дистантных связей в предикатных группах русской разговорной речи, а также дали представление о социолингвистической вариативности предмета исследования. Кроме того, настоящее исследование показало работоспособность избранных методов анализа устной речи. Также было выявлено несколько тенденций, статистическая значимость которых не была доказана.

По этой причине представляется возможным отнести к ***перспективам*** данного исследования дальнейшее изучение предикатных групп на большем количестве материала в двух направлениях. Во-первых, проведение более глубокого анализа дистантных связей с точки зрения не только количественных, но и качественных характеристик. Во-вторых, подтверждение или опровержение описанных наблюдений и формирование диагностических признаков социолингвистической вариативности устной разговорной речи.

# СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Аванесов Р.И. Русское литературное произношение: Уч. пособие. — М.: Просвещение, 1984. — 384 с.
2. Адмони В.Г. Система форм речевого высказывания. — СПб.: Наука, 1994. — 153 с.
3. Ананьев Б.Г. Человек как предмет познания. — Л.: ЛГУ, 1968. — 338 с.
4. Бахтин М.М. Вопросы литературы и эстетики. — М.: Художественная литература, 1975. — 504 с.
5. Бахтин М.М. Эстетика словесного творчества. — М.: Искусство, 1986. — 445 с.
6. Беликов В.И., Крысин Л.П. Социолингвистика. Учебник. — М.: РГГУ, 2001. — 439 с.
7. Богданов В.В. Семантико-синтаксическая организация предложения. — Л.: ЛГУ, 1977. — 203 с.
8. Богданова Н.В. О корпусе текстов живой речи: новые поступления и первые результаты исследования // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Междун. конфер. «Диалог» (Бекасово, 26-30 мая 2010 г.). — Вып. 9 (16). / гл. ред. А.Е. Кибрик. — М.: РГГУ, 2010а. — С. 35–40.
9. Богданова Н.В. Полевая лингвистика и проблема учебного текста // Материалы XXXVI Междун. филол. конф. 12-17 марта 2007 г., Санкт-Петербург. — Вып. 20: Полевая лингвистика. Интегральное моделирование звуковой формы естественного языка / отв. ред. А.С. Асиновский, Н.В. Богданова. — СПб.: СПбГУ, 2007. — С. 21–30.
10. Богданова Н.В. Принцип невода и принцип ковчега в корпусных исследованиях русской речи // Слово есть дело. Юбилейный сборник научных трудов в честь профессора Ирины Павловны Лысаковой. — Т. 1. — СПб.: Сударыня, 2010б. — С. 82–91.
11. Богданова Н.В., Бродт И.С., Куканова В.В., Павлова О.В., Сапунова Е.М., Филиппова Н.С. О «корпусе» текстов живой речи: принципы формирования и возможности описания // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной междун. конф. «Диалог» (2008). — Вып. 7 (14). / гл. ред. А.Е. Кибрик. — М.: РГГУ, 2008. — С. 57–61.
12. Богданова Н.В., Куканова В.В., Сапунова Е.М., Филиппова Н.С. Звуковой корпус как способ описания языка современного города // Материалы Междун. науч. конф. «Речевая коммуникация в современной России». 28-30 апреля 2009 г. / под ред. О.С. Иссерс, Н.А. Кузьминой. — Омск: Ом. гос. ун-т, 2009. — С. 43–48.
13. Богданова-Бегларян Н.В., Шерстинова Т.Ю., Баева Е.М., Блинова О.В., Мартыненко Г.Я., Ермолова О.Б., Рыко А.И. и др. Русский язык повседневного общения: особенности функционирования в разных социальных группах. Коллективная монография. — СПб.: Лайка, 2016. — 244 с.
14. Богданова-Бегларян Н.В. Прагматемы в устной повседневной речи: определение понятия и общая типология // Вестник Пермского университета. Российская и зарубежная филология. — Вып. 3 (27). — Пермь: ПГУ, 2014. — С. 7–20.
15. Богданова-Бегларян Н.В. Функционирование некоторых прагматем русской устной речи в коммуникации представителей разных социальных групп // Вестник Пермского университета. Российская и зарубежная филология. — Вып. 2 (34). — Пермь: ПГУ, 2016. — С. 38–49.
16. Богданова-Бегларян Н.В., Асиновский А.С., Блинова О.В., Маркасова Е.В., Рыко А.И., Шерстинова Т.Ю. Звуковой корпус русского языка: новая методология анализа устной речи // Язык и метод: Русский язык в лингвистических исследованиях XXI века. — Вып. 2 / ред. Д. Шумска, К. Озга. — Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 2015. — С. 357–372.
17. Богданова-Бегларян Н.В., Блинова О.В., Мартыненко Г.Я., Шерстинова Т.Ю. Некоторые инвариантные характеристики русской разговорной речи: фонетика, морфология, синтаксис // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной Междун. конф. «Диалог» (Москва, 31 мая — 3 июня 2017 г.). — Вып. 16 (23). Т. 2. — М.: РГГУ, 2017а. — С. 75–86.
18. Богданова-Бегларян Н.В., Шерстинова Т.Ю., Блинова О.В., Мартыненко Г.Я. Корпус «Один речевой день» в исследованиях социолингвистической вариативности русской разговорной речи // Анализ разговорной русской речи (АР3-2017): Труды седьмого междисциплинарного семинара / науч. ред. Д.А. Кочаров, П.А. Скрелин. — СПб.: Политехника-принт, 2017б. — С. 14–20.
19. Бодуэн де Куртенэ И.А. Значение языка как предмета изучения // Методика преподавания русского языка (хрестоматия) / ред. М.С. Лапатухин. — М.: Учпедгиз, 1960. — 483 с.
20. Бодуэн де Куртенэ И.А. Некоторые из общих положений, к которым довели Бодуэна его наблюдения и исследования явлений языка // И.А. Бодуэн де Куртенэ. Избранные труды по общему языкознанию. — Т. 1. — М.: АН СССР, 1963. — 384 с.
21. Брайт У. Введение: параметры социолингвистики // Новое в лингвистике. — Вып. VII. — М.: Изд-во иностранной литературы, 1975. — С. 34–41.
22. Буторов В.Д. Моделирование синтаксиса естественного языка // Прикладное языкознание / отв. ред. А.С. Герд. — СПб.: СПбГУ, 1996. — 528 с.
23. Вахтин Н.Б., Головко Е.В. Социолингвистика и социология языка. Учебное пособие. — СПб.: Гум. академия, 2004. — 336 с.
24. Венгер Л.А. Возрастная психология // Большая советская энциклопедия. — Т. 5. — М.: Советская энциклопедия, 1971. — 640 с.
25. Волохина Г.А., Попова З.Д. Синтаксические концепты русского простого предложения. — Воронеж: ВГУ, 2003. — 196 с.
26. Выготский Л.С. Мышление и речь // Л. С. Выготский. Избранные психологические исследования. — М.: АПН РСФСР, 1956. — 519 с.
27. Гаспаров Б.М. Устная речь как семиотический объект // Ученые записки Тартуского университета. — Вып. 442. — Тарту: Тартусский университет, 1978. — 145 с.
28. Головин Б.Н. Вопросы социальной дифференциации языка // Вопросы социальной лингвистики. — Л.: Наука, 1969. — 417 с.
29. Гомон Т.В. Исследование документов с деформированной внутренней структурой.: дис. … канд. юрид. наук. — М., 1990. — 182 с. (машинопись).
30. Горошко Е.И. Особенности мужского и женского вербального поведения (психолингвистический анализ): дис. … канд. филол. наук. — М., 1999. — 158 с. (машинопись).
31. Грановская Р.М. Элементы практической психологии: Уч. пособие. — Л.: ЛГУ, 1988. — 560 с.
32. Ельмслев Л. Пролегомены к теории языка // Новое в лингвистике. — Вып. 1. — М.: Изд-во иностранной литературы, 1960. — С. 264–389.
33. Ерофеева Е.В. Вероятностная структура идиомов: социолингвистический аспект. — Пермь: ПГУ, 2005. — 315 с.
34. Ерофеева Е.В. Языковая ситуация современного города и методы ее исследования // Вестник Пермского университета. Российская и зарубежная филология. — Вып. 5 — Пермь: ПГУ, 2009. — С. 16–23.
35. Ерофеева Т.И. Опыт исследования речи горожан (территориальный, социальный и психологический аспекты). — Свердловск: Урал. гос. ун-т, 1991. — 135 с.
36. Ерофеева Т.И. Понятие «социолект» в истории лингвистики XX века // Изменяющийся языковой мир. — Пермь: ПГУ, 2002. — С. 170–175.
37. Ерофеева Т.И. Современная городская речь. — Пермь: ПГУ, 2004. — 316 с.
38. Ерофеева Т.И. Штрихи речевого портрета Прикамья. Учебное пособие. — М.: Флинта, 2012. — 192 с.
39. Жинкин Н.И. Механизмы речи. — М.: АПН РСФСР, 1958. — 370 с.
40. Захаров В.П. Корпусная лингвистика: Уч.-метод. пособие. — СПб.: СПбГУ, 2005. — 48 с.
41. Звуковой корпус как материал для анализа русской речи. Коллективная монография. Часть 1. Чтение. Пересказ. Описание / отв. ред. Н.В. Богданова-Бегларян. — СПб.: СПбГУ, 2013. — 532 с.
42. Звуковой корпус как материал для анализа русской речи. Коллективная монография. Часть 2. Теоретические и практические аспекты анализа. Том 1. О некоторых особенностях устной спонтанной речи разного типа. Звуковой корпус как материал для преподавания русского языка в иностранной аудитории / отв. ред. Н.В. Богданова-Бегларян. — СПб.: СПбГУ, 2014. — 396 с.
43. Звуковой корпус как материал для анализа русской речи. Коллективная монография. Часть 2. Теоретические и практические аспекты анализа. Том 2. Звуковой корпус как материал для новых лексикографических проектов / отв. ред. Н.В. Богданова-Бегларян. — СПб.: СПбГУ, 2015. — 364 с.
44. Земская Е.А. Русская разговорная речь: лингвистический анализ и проблемы обучения. — М.: Русский язык, 1979. — 240 с.
45. Земская Е.А., Китайгородская М.А., Розанова Н.Н. Особенности мужской и женской речи // Русский язык в его функционировании. Коммуникативно-прагматический аспект. — М.: Наука, 1993. — С. 90–135.
46. Земская Е.А., Китайгородская М.В., Ширяев Е.Н. Русская разговорная речь. Общие вопросы. Словообразование. Синтаксис. — М.: Наука, 1981. — 276 с.
47. Иванова О.А. К характеристике внутриязыкового контакта между литературной и профессиональной речью носителя русского языка // Материалы XXXVII междун. филол. конф. — Вып. 21. Полевая лингвистика. Интегральное моделирование звуковой формы естественных языков. 10–15 марта 2008 года / отв. ред. А.С. Асиновский, науч. ред. Н.В. Богданова. — СПб.: СПбГУ, 2008. — С. 25–35.
48. Иванова О.А. Специфика бытовой речи различных профессионально ориентированных групп: дис. … маг. лингв. — СПб., 2011. — 112 с. (машинопись).
49. Иванова О.А. Стилевая «разноголосица» в бытовой спонтанной русской речи // Вестник Санкт-Петербургского университета. Филология. Востоковедение. Журналистика. Серия 9. — СПб.: СПбГУ, 2010. — С. 116–121.
50. Ильичева О.В. Синтаксические признаки спонтанной речи в соотношении с социальными характеристиками говорящего // Материалы XXVII межвуз. науч.-метод. конф. преподавателей и аспирантов. — Вып. 2. Секция грамматики (русско-славянский цикл). 16-22 марта 1998 г.: Тезисы докладов / отв. ред. Г.Н. Акимова. — СПб.: СПбГУ, 1998. — С. 32–35.
51. Карасик В.И. Язык социального статуса. — М.: ИЯ АН СССР, Волгоградский педагогический ин-т, 1991. — 495 с.
52. Кибрик А.А., Подлесская В.И. Рассказы о сновидениях. Корпусное исследование устного русского дискурса. — М.: Языки славянских культур, 2009. — 735 c.
53. Кибрик А.А. Полевая лингвистика // Энциклопедия Кругосвет: универсальная научно-популярная онлайн-энциклопедия. 2017. URL: http://www.krugosvet.ru/enc/gumanitarnye\_nauki/lingvistika/polevaya\_lingvistika.html (дата обращения: 19.05.2018).
54. Кирилина А.В., Томская М.В. Лингвистические гендерные исследования // Отечественные записки. — № 2 (23). — М.: НИУ ВШЭ, 2005. — С. 112–132.
55. Косериу Э. Синхрония, диахрония и история // Новое в лингвистике. — Вып. 3. — М.: Изд-во иностранной литературы, 1963. — С. 143–343.
56. Крысин Л.П. Социолингвистические аспекты изучения современного русского языка. — М.: Наука, 1989. — 186 с.
57. Крысин Л.П. Язык в современном обществе. — М.: Просвещение, 1977. — 191 с.
58. Ладыженская Т.А. Живое слово: Устная речь как средство и предмет обучения. Учебное пособие. — М.: Просвещение, 1986. — 127 с.
59. Ладыженская Т.А. Система работы по развитию связной устной речи учащихся. — М.: Просвещение, 1975. — 215 с.
60. Лаптева О.А. Русский разговорный синтаксис / отв. ред. Ф.П. Филин. — М.: Книжный дом «ЛИБРОКОМ», 2013. — 400 с.
61. Лаптева О.А. Синтаксис типизированных предикативных конструкций устно-разговорной разновидности современного русского литературного языка: автореф. дис. … докт филол. наук. — М., 1974а. — 47 с.
62. Лаптева О.А. Устно-литературная разновидность современного русского литературного языка и другие его компоненты. Статья 1. // Вопросы стилистики. — Вып. 7. — Саратов: Сарат. гос. ун-т., 1974б. — С. 73–91.
63. Лаптева О.А. Устно-литературная разновидность современного русского литературного языка и другие его компоненты. Статья 2. // Вопросы стилистики. — Вып. 8. — Саратов: Сарат. гос. ун-т., 1974в. — С. 86–107.
64. Лаптева О.А. Устно-литературная разновидность современного русского литературного языка и другие его компоненты. Статья 3. // Вопросы стилистики. — Вып. 9. — Саратов: Сарат. гос. ун-т., 1975. — С. 93–104.
65. Леонтьев А.А. Язык, речь, речевая деятельность. — М.: Просвещение, 1969. — 214 с.
66. Лингвистический энциклопедический словарь / под ред. В.Н. Ярцевой. — М.: Советская энциклопедия, 1990. — 682 с.
67. Лурия А.Р. Язык и сознание. — М.: МГУ, 1979. — 320 с.
68. Малиновская А.И. Репродуктив как объект многоаспектного анализа (на материале Звукового корпуса русского языка): дис. ... маг. лингв. — СПб, 2012. — 103 с. (машинопись).
69. Мартине А. Основы общей лингвистики // Новое в лингвистике. — Вып. 3. — М.: Изд-во иностранной литературы, 1963. — С. 366–566.
70. Мартыненко Г.Я. Синтаксис живой спонтанной речи: симметрия линейных порядков // Корпусная лингвистика. 2015. Труды международной конференции / отв. ред.: В.П. Захаров, О.А. Митрофанова, М.В. Хохлова. — СПб.: СПбГУ, 2015. — С. 371–378.
71. Мартыненко Г.Я. Сложность синтаксических структур и стилистическая диагностика // Прикладное языкознание / отв. ред. А.С. Герд. — СПб.: СПбГУ, 1996. — 528 с.
72. Мартыненко Г.Я., Реброва П.В., Шерстинова Т.Ю., Ядченко Я.Н. Динамическая симметрия в исследованиях устной речи и письменного текста // Мат-лы XLIV Межд. филолог. конференции (2015) / ред.: М.В. Хохлова. Секция прикладной и математической лингвистики. — СПб.: СПбГУ, 2016. — В печати.
73. Маслов Ю.С. Введение в языкознание. Учебник для филологических специальностей вузов. — М.: Высшая школа, 1987. — 272 с.
74. Остин Дж.Л. Слово как действие // Новое в зарубежной лингвистике. — Вып. 17. Теория речевых актов. — М.: Прогресс, 1986. — С. 22–131.
75. Панов М.В. История русского литературного произношения XVIII-XX вв. — М.: Наука, 1990. — 456 с.
76. Панов М.В. Современный русский язык. Фонетика. — М.: Высшая школа, 1979. — 256 с.
77. Петерсон М.Н. Очерк синтаксиса русского языка. — М., Пг.: Госиздат, 1923. — 132 с.
78. Полевая лингвистическая практика: Учебно-методический комплекс сложной структуры. Ч. 2: Методические указания по обработке, многоуровневой разметке и лингвистическому анализу корпуса звучащих текстов на русском языке / Сост. Асиновский А.С., Богданова Н.В., Бродт И.С., Королева И.В., Русакова М.В., Рыко А.И., Степанова С.Б., Филиппова Н.С., Шерстинова Т.Ю. — СПб.: СПбГУ, 2010. — 141 с.
79. Попова Т.И. Сложные прагматемы-редупликаты в русской устной речи: проблемы описания и варианты решения // Коммуникативные исследования / гл. ред. О.С. Иссерс. — № 3 (9). — Омск: ОГУ, 2016. — С. 36–47.
80. Реброва П.В. Структурные и линейные порядки в спонтанной речи (на материале корпуса «Один речевой день»): дис. ... маг. лингв. — СПб., 2014. — 104 с. (машинопись).
81. Реброва П.В. Формальное описание синтаксиса устной речи при помощи грамматики непосредственно составляющих (на материале корпуса «Один речевой день»): дис. … канд. филол. наук. — СПб., 2017. — 109 с. (машинопись).
82. Рубинштейн C.JI. Основы общей психологии. — М.: Учпедгиз, 1940. — 704 с.
83. Русакова Е.Б. Русский компьютерный социолект: формирование и функционирование: дис. … канд. филол. наук. — Калининград, 2007. — 192 с. (машинопись).
84. Русский язык повседневного общения: особенности функционирования в разных социальных группах. Коллективная монография / отв. ред. Н.В. Богданова-Бегларян. — СПб.: ЛАЙКА, 2016. — 244 с.
85. Салматова Ю.О. Статистический анализ глагольных групп в русской устной речи: дипломная работа. — СПб.: СПбГУ, 2010. — 40 с.
86. Седов К.Ф. Дискурс и личность эволюция коммуникативной компетенции. — М.: Лабиринт, 2004. — 320 с.
87. Серль Дж.Р. Что такое речевой акт? // Новое в зарубежной лингвистике. — Вып. 17. Теория речевых актов. — М.: Прогресс, 1986. — С. 151–169.
88. Сибата Т. Исследование языкового существования в течение 24 часов // Языкознание в Японии. — М.: Радуга, 1983. — С. 134–141.
89. Сиротинина О.Б. Разговорная речь (определение понятия, основные проблемы) // Вопросы социальной лингвистики. — Л.: Наука, 1969. — 373–391.
90. Сиротинина О.Б. Русская разговорная речь: Пособие для учителя — М.: Просвещение, 1983. — 80 с.
91. Словарь социолингвистических терминов / Отв. ред В.Ю. Михальченко. — М.: ИЯ РАН, 2006. — 312 с.
92. Смирницкий А.И. Синтаксис английского языка. — М.: Изд-во литературы на иностранных языках, 1957. — 286 с.
93. Соссюр Ф. де. Заметки по общей лингвистике. 2-е изд. — М.: Прогресс, 2001. — 280 с.
94. Соссюр Ф. де. Курс общей лингвистики. — Екатеринбург: Урал. гос. ун-т, 1999. — 432 с.
95. Социолингвистика // Энциклопедия Кругосвет. Универсальная научно-популярная онлайн-энциклопедия. 2017. URL: http://www.krugosvet.ru/enc/gumanitarnye\_nauki/sociologiya/sotsiolingvistika.html (дата обращения: 19.05.2018).
96. Социология // Энциклопедия / Сост. А. А. Грицанов, В. Л. Абушенко, Г. М. Евелькин, Г. Н. Соколова, О. В. Терещенко. — М.: Книжный Дом, 2003. — 1312 с.
97. Степанова С.Б. Темп речи как отражение социальных и психологических характеристик говорящего // Approaches to Slavic Interaction. University of Potsdam. Conference, March 16-18, 2011. — Potsdam, 2011. — С. 44–45.
98. Степанова С.Б., Асиновский А.С., Богданова Н.В., Русакова М.В. Шерстинова Т. Ю. Звуковой корпус русского языка повседневного общения «Один речевой день»: концепция и состояние функционирования // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной междун. конф. «Диалог» (2008). — Вып. 7 (14). / ред. А.Е. Кибрик. — М.: РГГУ, 2008. — С. 488–494.
99. Тарланов З.К. Избранные работы по языкознанию и филологии. — Петрозаводск: ПетрГУ, 2005. — 784 с.
100. Теньер Л. Основы структурного синтаксиса. Перевод с французского. — М.: Прогресс, 1988. — 656 с.
101. Тестелец Я.Г. Введение в общий синтаксис. — М.: РГГУ, 2001. — 798 с.
102. Трошева Т.Б. Письменная речь // Стилистический энциклопедический словарь русского языка / под ред. М.Н. Кожиной. — М.: Флинта, Наука, 2006а. — С. 285–287.
103. Трошева Т.Б. Устная речь // Стилистический энциклопедический словарь русского языка / под ред. М.Н. Кожиной. — М.: Флинта, Наука, 2006б. — С. 567–569.
104. Цейтлин С.Н. Язык и ребенок: Лингвистика детской речи: Учеб. пособие для студ. высш. учеб. заведений. — М.: Гуманит. изд. центр ВЛАДОС, 2000. — 240 с.
105. Шведова Н.Ю. Очерки по синтаксису русской разговорной речи. — М.: Наука, 1960. — 170 с.
106. Шепелёв А.С. Симметрийные свойства глагольного куста в русском и английском языках: дипломная работа. — СПб.: СПбГУ, 2009. — 21 с.
107. Шерстинова Т.Ю. Наиболее употребительные слова повседневной русской речи (в гендерном аспекте и в зависимости от условий коммуникации) [The Most Frequent Words in Everyday Spoken Russian (in the gender dimension and depending on communication settings)] // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 1-4 июня 2016 г.). — Вып. 15 (22). — М.: РГГУ, 2016. — C. 616–631.
108. Шерстинова Т.Ю. Синтаксис повседневной русской речи сквозь призму N-грамм анализа // Материалы Симпозиума «Русская грамматика: структурная организация языка и процессы языкового функционирования». СПбГУ, 23-26 мая 2018. — СПб.: СПбГУ. — В печати.
109. Шмелев Д.Н. О функционально-стилистической дифференциации языка // Русский язык в его функциональных разновидностях. — М.: Наука, 1977. — 168 с.
110. Шмид B. Нарратология. — М.: Языки славянской культуры, 2003. — 312 с.
111. Штерн А.С. Введение в психологию. Курс лекций. — М.: Флинта, МПСИ, 2006. — 162 с.
112. Щерба Л.В. О служебном и самостоятельном значении грамматики как учебного предмета // Л. В. Щерба. Избранные работы по русскому языку. — М.: Учпедгиз, 1957а. — 186 с.
113. Щерба Л.В. О трояком аспекте языковых явлений и об эксперименте в языкознании // Л. В. Щерба. Языковая система и речевая деятельность. — Л.: Наука, 1974. — 427 c.
114. Щерба Л.В. Современный русский литературный язык // Л. В. Щерба. Избранные работы по русскому языку. — М.: Учпедгиз, 1957б. — 186 с.
115. Эйсмонт П.М. Семантика и синтаксис спонтанного нарратива: экспериментальное исследование на материале русского языка: дис. … канд. филол. наук. — СПб., 2008. — 279 с. (машинопись).
116. Эйсмонт П.М. Стратегии порождения спонтанного нарратива и роль внешнего стимула // Вестник СПбГУ. Язык и литература. 2007. — Вып. 4. Ч. II. — С. 191–197.
117. Эриксон Э.Г. Детство и общество. — СПб.: Летний сад, 2000. — 415 с.
118. Якобсон Р.О. Лингвистика и поэтика // Структурализм: «за» и «против». — М.: Прогресс, 1975. — С. 193–230.
119. Якобсон Р.О. Язык и бессознательное. — М.: Гнозис, 1996. — 248 с.
120. Яскевич А.А. Корпусная лингвистика // Энциклопедия для школьников и студентов: в 12 т. Т. 1: Информационное общество. ХХI век / под общ. ред В. И. Стражева. — Минск: Белорусская энциклопедия, 2009. — С. 167–169.
121. Asinovsky А., Bogdanova N., Rusakova М., Ryko A., Stepanova S., Sherstinova T. The ORD Speech Corpus of Russian Everyday Communication «One Speaker's Day»: Creation Principles and Annotation / V. Matoušek, P. Mautner (eds.). // — LNAI, vol. 5729. — Springer, Berlin-Heidelberg, 2009. — Pp. 250–257.
122. Bogdanova-Beglarian N., Martynenko G. and Sherstinova T. The “One Day of Speech” Corpus: Phonetic and Syntactic Studies of Everyday Spoken Russian / Ronzhin, A. et al. (eds.) SPECOM 2015, Lecture Notes in Artificial Intelligence. — LNAI, vol. 9319. — Springer International Publishing Switzerland, 2015. —  Pp. 429–437.
123. Bogdanova-Beglarian N., Sherstinova T., Blinova O., Ermolova O., Baeva E., Martynenko G., Ryko A. Sociolinguistic Extension of the ORD Corpus of Russian Everyday Speech / Ronzhin, A.  et al. (eds.) SPECOM 2016, Lecture Notes in Artificial Intelligence. — LNAI, vol. 9811. — Springer, Switzerland, 2016. — Pp. 659–666.
124. Coupland N., Coupland J., Giles H. Language, Society, and the Elderly: Discourse, Identity, and Aging. — Oxford UK, Cambridge USA: Blackwell, 1991. — 220 p.
125. Eysmont P.M. Semantics vs. syntax in verb structure acquisition / Bertinetto P.M. et al. (eds.). Interdisciplinary Workshop on Verbs. The Identification and Representation of Verb Features. Proceedings of Verb. —Pisa: Scuola Normale Superiore, 2010. — Pp. 123–136.
126. Fishman P.M. Interaction: The Work Women Do // Social Problems. —Vol. 25, No. 4. — Oxford UK: Oxford University Press, 1978. — Pp. 397–406.
127. Hymes D.H. Foundations of Sociolinguistics: An Ethnographic Approach. — Philadelphia: University of Pennsylvania Press, 1974. — 245 p.
128. Labov W. The Social Stratification of English in New York city. — New York: Cambridge University Press, 2006. — 485 p.
129. Labov W. Variation in Language // The Learning of Language. — New York: Appleton-Century Crofts, 1971. — Pp. 187–221.
130. Martynenko G., Yadchenko Y. Quantitative language typology based on symmetry properties of syntactic structures // New Developments in the Quantitative Study of Languages (Book of abstracts). 28-29 August 2015. — Helsinki, 2015. — pp. 58–59.
131. Palek B. Cross-Reference. A Study from Hyper-Syntax. — Prague: Universita Karlova, 1968. — 221 p.
132. Ryan E. B. Why do Low-Prestige Language Varieties Persist? // Language and Social Psychology. — Oxford: Blackwell, 1979. — Pp. 145–157.
133. Sacks H., Schegloff E. Opening up Closings // Semiotica. — Vol. 8. No. 4. — Berlin: de Gruyter Mouton, 1973. — Pp. 289–327.
134. Sacks H., Schegloff E. A., Jefferson G. A Simplest Systematic for the Organization of Turn-Taking for Conversation // Language. — Vol. 50. No. 4, Part I. — Oxford: Oxford University Press, 1974. — Pp. 696–735.
135. Tannen D. You Just Don’t Understand: Women and Men in Conversation. — New York: William Morrow & Co., 1990. — 330 p.
136. Trudgill P. Sociolinguistics: An Introduction to Language and Society. — London: Penguin Books, 1983. — 109 p.
137. Zimin S. Sex and Politeness: Factors in First and Second Language Use // International Journal of the Sociology of Language. — Vol. 27. — Berlin: de Gruyter Mouton, 1981. — Pp. 35–58.

# СПИСОК СОКРАЩЕНИЙ, ПРИНЯТЫХ В РАБОТЕ

ГУМ – представители гуманитарной сферы

ЗКРЯ – Звуковой корпус русского языка

КЛЯ – кодифицированный литературный язык

КУТ – корпус устных текстов

ЛЯ – литературный язык

ОБР – работники сферы образования

ОРД – Один речевой день

ОФ – офисные работники

ПГ – предикатная группа

РА – речевой акт

РР – разговорная речь

СИЛ – представители силовых структур

ТВОР – представители творческих профессий

УР – устная речь

ЭДЕ – элементарная дискурсивная единица

IT – IT-специалисты

# ПРИЛОЖЕНИЯ

## Приложение. Таблицы

*Таблица 9*

*Описательная статистика предикатных групп устной спонтанной речи информантов-мужчин*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 332 | 4.289157 | 1.000000 | 17.00000 | 2.394758 |
| Левая дистантность | 332 | 2.439759 | 0.000000 | 11.00000 | 1.922828 |
| Правая дистантность | 332 | 0.849398 | 0.000000 | 10.00000 | 1.557042 |

*Таблица 10*

*Описательная статистика предикатных групп устной спонтанной речи информантов-женщин*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 498 | 4.275100 | 1.000000 | 14.00000 | 2.116958 |
| Левая дистантность | 498 | 2.520080 | 0.000000 | 10.00000 | 1.810043 |
| Правая дистантность | 498 | 0.761044 | 0.000000 | 7.00000 | 1.199064 |

*Таблица 11*

*Описательная статистика предикатных групп устной спонтанной речи информантов до 30 лет*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 263 | 4.665399 | 1.000000 | 17.00000 | 2.383591 |
| Левая дистантность | 263 | 2.863118 | 0.000000 | 10.00000 | 1.907270 |
| Правая дистантность | 263 | 0.802281 | 0.000000 | 10.00000 | 1.411132 |

*Таблица 12*

*Описательная статистика предикатных групп устной спонтанной речи информантов от 31 года до 45 лет*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 339 | 3.941003 | 1.000000 | 13.00000 | 2.104391 |
| Левая дистантность | 339 | 2.129794 | 0.000000 | 9.00000 | 1.754360 |
| Правая дистантность | 339 | 0.814159 | 0.000000 | 9.00000 | 1.263146 |

*Таблица 13*

*Описательная статистика предикатных групп устной спонтанной речи информантов от 46 лет*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 228 | 4.342105 | 1.000000 | 15.00000 | 2.162267 |
| Левая дистантность | 228 | 2.587719 | 0.000000 | 11.00000 | 1.850403 |
| Правая дистантность | 228 | 0.763158 | 0.000000 | 9.00000 | 1.419205 |

*Таблица 14*

*Описательная статистика предикатных групп устной спонтанной речи представителей гуманитарных наук*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 98 | 4.755102 | 1.000000 | 12.00000 | 2.364084 |
| Левая дистантность | 98 | 3.061224 | 0.000000 | 10.00000 | 2.089820 |
| Правая дистантность | 98 | 0.693878 | 0.000000 | 5.00000 | 1.134357 |

*Таблица 15*

*Описательная статистика предикатных групп устной спонтанной речи работников сферы образования*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 244 | 4.692623 | 1.000000 | 15.00000 | 2.251608 |
| Левая дистантность | 244 | 2.790984 | 0.000000 | 11.00000 | 1.759560 |
| Правая дистантность | 244 | 0.901639 | 0.000000 | 9.00000 | 1.549777 |

*Таблица 16*

*Описательная статистика предикатных групп устной спонтанной речи IT-специалистов*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 153 | 4.071895 | 1.000000 | 10.00000 | 1.871196 |
| Левая дистантность | 153 | 2.477124 | 0.000000 | 8.00000 | 1.792106 |
| Правая дистантность | 153 | 0.607843 | 0.000000 | 4.00000 | 1.001547 |

*Таблица 17*

*Описательная статистика предикатных групп устной спонтанной речи офисных работников*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 198 | 4.035354 | 1.000000 | 17.00000 | 2.410586 |
| Левая дистантность | 198 | 2.101010 | 0.000000 | 9.00000 | 1.846907 |
| Правая дистантность | 198 | 0.939394 | 0.000000 | 10.00000 | 1.489852 |

*Таблица 18*

*Описательная статистика предикатных групп устной спонтанной речи представителей силовых структур*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 49 | 3.979592 | 1.000000 | 8.000000 | 1.725903 |
| Левая дистантность | 49 | 2.510204 | 0.000000 | 6.000000 | 1.721463 |
| Правая дистантность | 49 | 0.469388 | 0.000000 | 2.000000 | 0.738863 |

*Таблица 19*

*Описательная статистика предикатных групп устной спонтанной речи представителей творческих профессий*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Количество анализируемых единиц | Среднее значение | Минимальное значение | Максимальное значение | Стандартное отклонение |
| Размер ПГ | 88 | 3.693182 | 1.000000 | 13.00000 | 2.183312 |
| Левая дистантность | 88 | 1.886364 | 0.000000 | 8.00000 | 1.724948 |
| Правая дистантность | 88 | 0.806818 | 0.000000 | 9.00000 | 1.421170 |

*Таблица 20*

*Ранжированный список частеречных структур ПГ*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № | структура | количество в анализируемом материале | доля в % | ранг |
| 1 | *V* | 45 | 5.42 | 1 |
| 2 | *S-PRO V* | 17 | 2.05 | 2 |
| 3 | *PART V* | 16 | 1.93 | 3 |
| 4 | *CONJ S-PRO V* | 15 | 1.81 | 4 |
| 5 | *PRAEDIC* | 13 | 1.57 | 5 |
| 6 | *S V* | 12 | 1.45 | 6 |
| 7 | *CONJ V* | 10 | 1.20 | 7 |
| 8 | *ADV-PRO V* | 10 | 1.20 | 7 |
| 9 | *V S* | 10 | 1.20 | 7 |
| 10 | *PART PART V* | 9 | 1.08 | 8 |
| 11 | *PART S-PRO V* | 7 | 0.84 | 9 |
| 12 | *S-PRO PART V* | 7 | 0.84 | 9 |
| 13 | *PART S-PRO PART V* | 6 | 0.72 | 10 |
| 14 | *V V* | 6 | 0.72 | 10 |
| 15 | *V S-PRO* | 4 | 0.48 | 11 |
| 16 | *PRAEDIC V* | 3 | 0.36 | 12 |
| 17 | *CONJ S-PRO PART V* | 3 | 0.36 | 12 |
| 18 | *S-PRO V V* | 3 | 0.36 | 12 |
| 19 | *PART V S-PRO* | 3 | 0.36 | 12 |
| 20 | *PR S V* | 3 | 0.36 | 12 |
| 21 | *ADV-PRO S-PRO V* | 3 | 0.36 | 12 |
| 22 | *S-PRO ADV V* | 3 | 0.36 | 12 |
| 23 | *ADV PART V* | 3 | 0.36 | 12 |
| 24 | *V ADV* | 3 | 0.36 | 12 |
| 25 | *ADV* | 3 | 0.36 | 12 |
| 26 | *V PR S* | 3 | 0.36 | 12 |
| 27 | *S-PRO PART PART V* | 3 | 0.36 | 12 |
| 28 | *S-PRO ADV-PRO V* | 3 | 0.36 | 12 |
| 29 | *S-PRO PR S V* | 3 | 0.36 | 12 |
| 30 | *CONJ V S* | 2 | 0.24 | 13 |
| 31 | *ADV-PRO PART V* | 2 | 0.24 | 13 |
| 32 | *A-PRO S PART V* | 2 | 0.24 | 13 |
| 33 | *ADV-PRO V PR S* | 2 | 0.24 | 13 |
| 34 | *ADV A* | 2 | 0.24 | 13 |
| 35 | *ADV-PRO PRAEDIC* | 2 | 0.24 | 13 |
| 36 | *ADV PRAEDIC* | 2 | 0.24 | 13 |
| 37 | *S-PRO V PR S* | 2 | 0.24 | 13 |
| 38 | *INTJ S V* | 2 | 0.24 | 13 |
| 39 | *S-PRO A* | 2 | 0.24 | 13 |
| 40 | *CONJ S-PRO PART V S* | 2 | 0.24 | 13 |
| 41 | *S-PRO V ADV* | 2 | 0.24 | 13 |
| 42 | *PART S V* | 2 | 0.24 | 13 |
| 43 | *CONJ S-PRO V ADV* | 2 | 0.24 | 13 |
| 44 | *PART S-PRO PART PART V* | 2 | 0.24 | 13 |
| 45 | *CONJ V A S* | 2 | 0.24 | 13 |
| 46 | *PART S-PRO S-PRO V* | 2 | 0.24 | 13 |
| 47 | *CONJ S-PRO ADV-PRO V* | 2 | 0.24 | 13 |
| 48 | *S-PRO CONJ PART V* | 2 | 0.24 | 13 |
| 49 | *V PART* | 2 | 0.24 | 13 |
| 50 | *S-PRO PART V S* | 2 | 0.24 | 13 |
| 51 | *S-PRO S-PRO V* | 2 | 0.24 | 13 |
| 52 | *A-PRO S V* | 2 | 0.24 | 13 |
| 53 | *S-PRO V PART* | 2 | 0.24 | 13 |
| 54 | *PR S V ADV* | 2 | 0.24 | 13 |
| 55 | *S-PRO V S* | 2 | 0.24 | 13 |
| 56 | *CONJ ADV-PRO V* | 2 | 0.24 | 13 |
| 57 | *CONJ S-PRO V PR A-PRO* | 2 | 0.24 | 13 |
| 58 | *PRAEDIC S V* | 2 | 0.24 | 13 |
| 59 | *CONJ S V* | 2 | 0.24 | 13 |
| 60 | *PART V S* | 2 | 0.24 | 13 |
| 61 | *V S-PRO V PART* | 2 | 0.24 | 13 |
| 62 | *V A* | 1 | 0.12 | 14 |
| 63 | *S ADV-PRO A* | 1 | 0.12 | 14 |
| 64 | *PART V ADV PART* | 1 | 0.12 | 14 |
| 65 | *ADV-PRO S-PRO ADV-PRO* | 1 | 0.12 | 14 |
| 66 | *S-PRO PARENTH PR S-PRO CONJ V* | 1 | 0.12 | 14 |
| 67 | *ADV-PRO S-PRO A-PRO A* | 1 | 0.12 | 14 |
| 68 | *PART S-PRO ADV S-PRO A* | 1 | 0.12 | 14 |
| 69 | *ADV-PRO S-PRO CONJ V* | 1 | 0.12 | 14 |
| 70 | *PR S-PRO ANUM ADV V* | 1 | 0.12 | 14 |
| 71 | *ADV-PRO S-PRO PART PART V* | 1 | 0.12 | 14 |
| 72 | *S V PR S* | 1 | 0.12 | 14 |
| 73 | *ADV-PRO S-PRO PART S-PRO S V* | 1 | 0.12 | 14 |
| 74 | *S-PRO PRAEDIC* | 1 | 0.12 | 14 |
| 75 | *ADV-PRO S-PRO PR S V* | 1 | 0.12 | 14 |
| 76 | *PART PR S S-PRO ADV V* | 1 | 0.12 | 14 |
| 77 | *ADV-PRO S-PRO PRAEDIC* | 1 | 0.12 | 14 |
| 78 | *A S PART S-PRO V V* | 1 | 0.12 | 14 |
| 79 | *ADV-PRO S-PRO S-PRO ADV V A* | 1 | 0.12 | 14 |
| 80 | *PR S CONJ S-PRO S-PRO V V PART* | 1 | 0.12 | 14 |
| 81 | *ADV S-PRO ADV S-PRO PART V* | 1 | 0.12 | 14 |
| 82 | *ADV-PRO ADV-PRO V* | 1 | 0.12 | 14 |
| 83 | *ADV-PRO S-PRO V A-PRO S* | 1 | 0.12 | 14 |
| 84 | *S PART V ADV* | 1 | 0.12 | 14 |
| 85 | *ADV-PRO S-PRO V CONJ* | 1 | 0.12 | 14 |
| 86 | *S-PRO ADV-PRO ADV V* | 1 | 0.12 | 14 |
| 87 | *S-PRO PART S-PRO ANUM V* | 1 | 0.12 | 14 |
| 88 | *ADV-PRO S-PRO V PART PART* | 1 | 0.12 | 14 |
| 89 | *S-PRO V A* | 1 | 0.12 | 14 |
| 90 | *ADV-PRO S-PRO V S PR ADV-PRO S* | 1 | 0.12 | 14 |
| 91 | *V PART ADV S* | 1 | 0.12 | 14 |
| 92 | *A PARENTH PART PART V ADV* | 1 | 0.12 | 14 |
| 93 | *PART S NUM PART V* | 1 | 0.12 | 14 |
| 94 | *ADV-PRO V ADV* | 1 | 0.12 | 14 |
| 95 | *PART S-PRO ADV-PRO V S PR NUM* | 1 | 0.12 | 14 |
| 96 | *ADV-PRO V INTJ* | 1 | 0.12 | 14 |
| 97 | *ADV-PRO ADV-PRO PART S S-PRO PART V V* | 1 | 0.12 | 14 |
| 98 | *PART V S-PRO V ADV* | 1 | 0.12 | 14 |
| 99 | *ADV S-PRO ADV V ADV-PRO* | 1 | 0.12 | 14 |
| 100 | *PR S S-PRO PART PART V* | 1 | 0.12 | 14 |
| 101 | *A-PRO A ADV* | 1 | 0.12 | 14 |
| 102 | *PR S-PRO PRAEDIC S* | 1 | 0.12 | 14 |
| 103 | *A-PRO ADV PART ADV* | 1 | 0.12 | 14 |
| 104 | *ADV-PRO PART NUM PARENTH S V PR S* | 1 | 0.12 | 14 |
| 105 | *A-PRO ADV PRAEDIC*  | 1 | 0.12 | 14 |
| 106 | *S PARENTH PR NUM V* | 1 | 0.12 | 14 |
| 107 | *A-PRO PARENTH CONJ V* | 1 | 0.12 | 14 |
| 108 | *S S-PRO S S V* | 1 | 0.12 | 14 |
| 109 | *A-PRO PARENTH S-PRO V* | 1 | 0.12 | 14 |
| 110 | *S-PRO ADV PARENTH S-PRO V PR A-PRO S* | 1 | 0.12 | 14 |
| 111 | *A-PRO PART S-PRO V* | 1 | 0.12 | 14 |
| 112 | *S-PRO A-PRO V CONJ A-PRO S* | 1 | 0.12 | 14 |
| 113 | *A-PRO PR ANUM V* | 1 | 0.12 | 14 |
| 114 | *S-PRO PART PART A S* | 1 | 0.12 | 14 |
| 115 | *A-PRO PR S S V S* | 1 | 0.12 | 14 |
| 116 | *S-PRO PR S PART V PR A-PRO S* | 1 | 0.12 | 14 |
| 117 | *A-PRO S ADV-PRO V PR S-PRO A* | 1 | 0.12 | 14 |
| 118 | *S-PRO S-PRO PART ADV V* | 1 | 0.12 | 14 |
| 119 | *ADV S-PRO S-PRO A-PRO PART PRAEDIC* | 1 | 0.12 | 14 |
| 120 | *S-PRO V PR S S PART* | 1 | 0.12 | 14 |
| 121 | *A-PRO S PART V CONJ* | 1 | 0.12 | 14 |
| 122 | *V A-PRO* | 1 | 0.12 | 14 |
| 123 | *ADV S-PRO S-PRO S-PRO V* | 1 | 0.12 | 14 |
| 124 | *ADV-PRO S V A-PRO A-PRO* | 1 | 0.12 | 14 |
| 125 | *A-PRO S-PRO PART V* | 1 | 0.12 | 14 |
| 126 | *PART PRAEDIC S* | 1 | 0.12 | 14 |
| 127 | *A-PRO V ADV S* | 1 | 0.12 | 14 |
| 128 | *PART S PRAEDIC* | 1 | 0.12 | 14 |
| 129 | *A-PRO V PR S* | 1 | 0.12 | 14 |
| 130 | *PART S-PRO ADV-PRO*  | 1 | 0.12 | 14 |
| 131 | *A-PRO V V* | 1 | 0.12 | 14 |
| 132 | *PART S-PRO PART A-PRO PART S* | 1 | 0.12 | 14 |
| 133 | *CONJ A S V* | 1 | 0.12 | 14 |
| 134 | *PART S-PRO S* | 1 | 0.12 | 14 |
| 135 | *CONJ A S V V* | 1 | 0.12 | 14 |
| 136 | *PART S-PRO V PR NUM S* | 1 | 0.12 | 14 |
| 137 | *CONJ A V* | 1 | 0.12 | 14 |
| 138 | *PART V PR S-PRO PRAEDIC S PR S* | 1 | 0.12 | 14 |
| 139 | *CONJ ADV* | 1 | 0.12 | 14 |
| 140 | *PR A-PRO S CONJ PART ADV-PRO PART V* | 1 | 0.12 | 14 |
| 141 | *CONJ ADV ADV V* | 1 | 0.12 | 14 |
| 142 | *PR S PR S V* | 1 | 0.12 | 14 |
| 143 | *CONJ ADV ADV-PRO PART A V S* | 1 | 0.12 | 14 |
| 144 | *PR S-PRO ADV V PART NUM S* | 1 | 0.12 | 14 |
| 145 | *CONJ ADV ADV-PRO V PRAEDIC* | 1 | 0.12 | 14 |
| 146 | *PR S-PRO PART PR S PART V S* | 1 | 0.12 | 14 |
| 147 | *CONJ ADV PART S-PRO S-PRO V PRAEDIC* | 1 | 0.12 | 14 |
| 148 | *PR S-PRO V S* | 1 | 0.12 | 14 |
| 149 | *CONJ ADV PART V* | 1 | 0.12 | 14 |
| 150 | *ADV-PRO PART A-PRO S V* | 1 | 0.12 | 14 |
| 151 | *CONJ ADV PRAEDIC S ADV-PRO PART V* | 1 | 0.12 | 14 |
| 152 | *S ADV ADV V* | 1 | 0.12 | 14 |
| 153 | *CONJ ADV S PART V* | 1 | 0.12 | 14 |
| 154 | *S ADV-PRO V S* | 1 | 0.12 | 14 |
| 155 | *CONJ ADV S V V* | 1 | 0.12 | 14 |
| 156 | *S PART PART V* | 1 | 0.12 | 14 |
| 157 | *CONJ ADV S-PRO V* | 1 | 0.12 | 14 |
| 158 | *S S S-PRO S-PRO V* | 1 | 0.12 | 14 |
| 159 | *CONJ ADV V* | 1 | 0.12 | 14 |
| 160 | *ADV A S* | 1 | 0.12 | 14 |
| 161 | *CONJ ADV V NUM PR S S* | 1 | 0.12 | 14 |
| 162 | *CONJ ADV V PRAEDIC* | 1 | 0.12 | 14 |
| 163 | *S-PRO ADV PR S* | 1 | 0.12 | 14 |
| 164 | *CONJ ADV V S-PRO V S-PRO* | 1 | 0.12 | 14 |
| 165 | *ADV-PRO PART S-PRO S PART V* | 1 | 0.12 | 14 |
| 166 | *CONJ ADV-PRO ADV V PR S A-PRO S A PR S-PRO PR S* | 1 | 0.12 | 14 |
| 167 | *S-PRO CONJ S-PRO V* | 1 | 0.12 | 14 |
| 168 | *CONJ ADV-PRO PART NUM ANUM PR S PR S V S S* | 1 | 0.12 | 14 |
| 169 | *S-PRO PART A-PRO PART PR A-PRO NUM S V* | 1 | 0.12 | 14 |
| 170 | *CONJ ADV-PRO PRAEDIC* | 1 | 0.12 | 14 |
| 171 | *S-PRO PART PR NUM V* | 1 | 0.12 | 14 |
| 172 | *CONJ ADV-PRO PRAEDIC V* | 1 | 0.12 | 14 |
| 173 | *ADV-PRO PART V A-PRO A* | 1 | 0.12 | 14 |
| 174 | *CONJ ADV-PRO S PART PART PRAEDIC* | 1 | 0.12 | 14 |
| 175 | *S-PRO PR S-PRO ADV ADV V S* | 1 | 0.12 | 14 |
| 176 | *CONJ ADV-PRO S PART V* | 1 | 0.12 | 14 |
| 177 | *S-PRO S V* | 1 | 0.12 | 14 |
| 178 | *CONJ ADV-PRO S V* | 1 | 0.12 | 14 |
| 179 | *S-PRO S-PRO PART V S* | 1 | 0.12 | 14 |
| 180 | *CONJ ADV-PRO S-PRO ADV-PRO V* | 1 | 0.12 | 14 |
| 181 | *ADV-PRO PR S-PRO PART V S-PRO V* | 1 | 0.12 | 14 |
| 182 | *CONJ ADV-PRO S-PRO NUM V* | 1 | 0.12 | 14 |
| 183 | *ADV-PRO PRAEDIC A-PRO S PR S-PRO* | 1 | 0.12 | 14 |
| 184 | *CONJ ADV-PRO S-PRO V* | 1 | 0.12 | 14 |
| 185 | *V ADV-PRO A-PRO A-PRO S* | 1 | 0.12 | 14 |
| 186 | *ADV S-PRO V PART PR A S* | 1 | 0.12 | 14 |
| 187 | *ADV-PRO PRAEDIC V V* | 1 | 0.12 | 14 |
| 188 | *CONJ A-PRO A S-PRO V* | 1 | 0.12 | 14 |
| 189 | *V PR ANUM ANUM* | 1 | 0.12 | 14 |
| 190 | *CONJ A-PRO S ADV-PRO V* | 1 | 0.12 | 14 |
| 191 | *PART PR S NUM ADV-PRO PRAEDIC* | 1 | 0.12 | 14 |
| 192 | *CONJ A-PRO V PART* | 1 | 0.12 | 14 |
| 193 | *PART PR S-PRO V PART* | 1 | 0.12 | 14 |
| 194 | *CONJ A-PRO V S CONJ A-PRO S PART* | 1 | 0.12 | 14 |
| 195 | *PART S ADV ADV V A-PRO S* | 1 | 0.12 | 14 |
| 196 | *CONJ CONJ ADV-PRO S CONJ ADV-PRO NUM* | 1 | 0.12 | 14 |
| 197 | *PART S PART V* | 1 | 0.12 | 14 |
| 198 | *CONJ CONJ PR S S-PRO PART V S V* | 1 | 0.12 | 14 |
| 199 | *PART S-PRO A* | 1 | 0.12 | 14 |
| 200 | *CONJ CONJ S-PRO V* | 1 | 0.12 | 14 |
| 201 | *PART S-PRO ADV V* | 1 | 0.12 | 14 |
| 202 | *CONJ CONJ S-PRO V V* | 1 | 0.12 | 14 |
| 203 | *PART S-PRO ADV-PRO A-PRO* | 1 | 0.12 | 14 |
| 204 | *CONJ NUM S ADV* | 1 | 0.12 | 14 |
| 205 | *PART S-PRO PARENTH A* | 1 | 0.12 | 14 |
| 206 | *CONJ PARENTH S-PRO V* | 1 | 0.12 | 14 |
| 207 | *PART S-PRO PART S* | 1 | 0.12 | 14 |
| 208 | *CONJ PARENTH V* | 1 | 0.12 | 14 |
| 209 | *PART S-PRO PR A S V* | 1 | 0.12 | 14 |
| 210 | *CONJ PART ADV PART V* | 1 | 0.12 | 14 |
| 211 | *PART S-PRO S-PRO PR S PARENTH* | 1 | 0.12 | 14 |
| 212 | *CONJ PART ANUM PR S-PRO A PR S V PRAEDIC* | 1 | 0.12 | 14 |
| 213 | *PART S-PRO V ADV PART V* | 1 | 0.12 | 14 |
| 214 | *CONJ PART PART S V* | 1 | 0.12 | 14 |
| 215 | *PART S-PRO V V* | 1 | 0.12 | 14 |
| 216 | *CONJ PART PR S-PRO ADV-PRO CONJ V PART PR A-PRO SADV-PRO CONJ S-PRO PR* | 1 | 0.12 | 14 |
| 217 | *PART V PARENTH S* | 1 | 0.12 | 14 |
| 218 | *CONJ PART S-PRO PART S-PRO PART ADV PART V* | 1 | 0.12 | 14 |
| 219 | *ADV-PRO CONJ PR S V* | 1 | 0.12 | 14 |
| 220 | *CONJ PART S-PRO S-PRO ADV V* | 1 | 0.12 | 14 |
| 221 | *POS* | 1 | 0.12 | 14 |
| 222 | *CONJ PART V* | 1 | 0.12 | 14 |
| 223 | *PR A-PRO V S* | 1 | 0.12 | 14 |
| 224 | *CONJ PART V A S* | 1 | 0.12 | 14 |
| 225 | *PR S PART V* | 1 | 0.12 | 14 |
| 226 | *CONJ PART V S* | 1 | 0.12 | 14 |
| 227 | *PR S PRAEDIC* | 1 | 0.12 | 14 |
| 228 | *CONJ PART V S-PRO A-PRO S PART* | 1 | 0.12 | 14 |
| 229 | *ADV-PRO CONJ S-PRO PART V* | 1 | 0.12 | 14 |
| 230 | *CONJ PR A S ADV-PRO V* | 1 | 0.12 | 14 |
| 231 | *PR S-PRO ADV-PRO A-PRO S V* | 1 | 0.12 | 14 |
| 232 | *CONJ PR ANUM S ADV S-PRO V S-PRO* | 1 | 0.12 | 14 |
| 233 | *PR S-PRO PART A-PRO CONJ S* | 1 | 0.12 | 14 |
| 234 | *CONJ PR A-PRO PART V A* | 1 | 0.12 | 14 |
| 235 | *PR S-PRO PRAEDIC A S* | 1 | 0.12 | 14 |
| 236 | *CONJ PR NUM S-PRO ADV V* | 1 | 0.12 | 14 |
| 237 | *PR S-PRO S PRAEDIC* | 1 | 0.12 | 14 |
| 238 | *CONJ PR S S V A-PRO PART* | 1 | 0.12 | 14 |
| 239 | *CONJ PR S S-PRO V* | 1 | 0.12 | 14 |
| 240 | *PRAEDIC PART* | 1 | 0.12 | 14 |
| 241 | *CONJ PR S V S A* | 1 | 0.12 | 14 |
| 242 | *PRAEDIC S-PRO PR S CONJ PRAEDIC* | 1 | 0.12 | 14 |
| 243 | *CONJ PR S V V* | 1 | 0.12 | 14 |
| 244 | *PRAEDIC V ADV-PRO S PART* | 1 | 0.12 | 14 |
| 245 | *CONJ PR S-PRO S PART V* | 1 | 0.12 | 14 |
| 246 | *S ADV V* | 1 | 0.12 | 14 |
| 247 | *CONJ PRAEDIC ADV PART* | 1 | 0.12 | 14 |
| 248 | *S ADV-PRO S PARENTH V* | 1 | 0.12 | 14 |
| 249 | *CONJ PRAEDIC ADV S-PRO PART V S ADV* | 1 | 0.12 | 14 |
| 250 | *S CONJ PART V* | 1 | 0.12 | 14 |
| 251 | *CONJ PRAEDIC PART V A S V* | 1 | 0.12 | 14 |
| 252 | *S PART ADV* | 1 | 0.12 | 14 |
| 253 | *CONJ PRAEDIC PR S-PRO V* | 1 | 0.12 | 14 |
| 254 | *S PART S V* | 1 | 0.12 | 14 |
| 255 | *CONJ PRAEDIC V* | 1 | 0.12 | 14 |
| 256 | *S PRAEDIC* | 1 | 0.12 | 14 |
| 257 | *CONJ S ADV* | 1 | 0.12 | 14 |
| 258 | *S S-PRO ADV V* | 1 | 0.12 | 14 |
| 259 | *CONJ S ADV PR S* | 1 | 0.12 | 14 |
| 260 | *S S-PRO V CONJ ADV-PRO V* | 1 | 0.12 | 14 |
| 261 | *CONJ S ADV-PRO V* | 1 | 0.12 | 14 |
| 262 | *S V ADV-PRO* | 1 | 0.12 | 14 |
| 263 | *CONJ S PR A S V* | 1 | 0.12 | 14 |
| 264 | *ADV-PRO PART PART V* | 1 | 0.12 | 14 |
| 265 | *CONJ S PRAEDIC* | 1 | 0.12 | 14 |
| 266 | *S-PRO A S V ADV* | 1 | 0.12 | 14 |
| 267 | *ADV S-PRO V PR S A S* | 1 | 0.12 | 14 |
| 268 | *S-PRO ADV PART V PR A-PRO S* | 1 | 0.12 | 14 |
| 269 | *S-PRO ADV V PR S-PRO PART* | 1 | 0.12 | 14 |
| 270 | *CONJ S-PRO A A S* | 1 | 0.12 | 14 |
| 271 | *S-PRO ADV-PRO PR S ADV V* | 1 | 0.12 | 14 |
| 272 | *CONJ S-PRO A PR S* | 1 | 0.12 | 14 |
| 273 | *S-PRO A-PRO A* | 1 | 0.12 | 14 |
| 274 | *CONJ S-PRO ADV ADV V* | 1 | 0.12 | 14 |
| 275 | *ADV-PRO PART S-PRO S-PRO V* | 1 | 0.12 | 14 |
| 276 | *CONJ S-PRO ADV ADV-PRO* | 1 | 0.12 | 14 |
| 277 | *S-PRO CONJ V PR S* | 1 | 0.12 | 14 |
| 278 | *CONJ S-PRO ADV PR S V* | 1 | 0.12 | 14 |
| 279 | *S-PRO PART ADV-PRO A V* | 1 | 0.12 | 14 |
| 280 | *CONJ S-PRO ADV V* | 1 | 0.12 | 14 |
| 281 | *S-PRO PART CONJ PART PART V* | 1 | 0.12 | 14 |
| 282 | *CONJ S-PRO ADV V A-PRO S* | 1 | 0.12 | 14 |
| 283 | *S-PRO PART PART V ADV-PRO* | 1 | 0.12 | 14 |
| 284 | *CONJ S-PRO ADV-PRO PART A-PRO S ADV PART V* | 1 | 0.12 | 14 |
| 285 | *S-PRO PART PRAEDIC* | 1 | 0.12 | 14 |
| 286 | *CONJ S-PRO ADV-PRO PR ADV ADV V* | 1 | 0.12 | 14 |
| 287 | *S-PRO PART V ADV PART V* | 1 | 0.12 | 14 |
| 288 | *CONJ S-PRO ADV-PRO PR A-PRO S* | 1 | 0.12 | 14 |
| 289 | *S-PRO PR A-PRO S V* | 1 | 0.12 | 14 |
| 290 | *ADV V* | 1 | 0.12 | 14 |
| 291 | *S-PRO PR S V ADV* | 1 | 0.12 | 14 |
| 292 | *CONJ S-PRO A-PRO V* | 1 | 0.12 | 14 |
| 293 | *S-PRO PR S-PRO ADV-PRO PR S V* | 1 | 0.12 | 14 |
| 294 | *CONJ S-PRO CONJ PR S-PRO ADV-PRO V* | 1 | 0.12 | 14 |
| 295 | *S-PRO S A-PRO A S A-PRO PART PRAEDIC* | 1 | 0.12 | 14 |
| 296 | *CONJ S-PRO PART ADV-PRO V* | 1 | 0.12 | 14 |
| 297 | *S-PRO S-PRO ADV-PRO V* | 1 | 0.12 | 14 |
| 298 | *CONJ S-PRO PART ANUM S V ADV-PRO* | 1 | 0.12 | 14 |
| 299 | *S-PRO S-PRO PART V PART* | 1 | 0.12 | 14 |
| 300 | *CONJ S-PRO PART PARENTH PART V A-PRO S S* | 1 | 0.12 | 14 |
| 301 | *S-PRO S-PRO V PR A-PRO S* | 1 | 0.12 | 14 |
| 302 | *CONJ S-PRO PART PART ADV-PRO PR S V ADV-PRO* | 1 | 0.12 | 14 |
| 303 | *S-PRO V ADV-PRO* | 1 | 0.12 | 14 |
| 304 | *ADV V ADV PRAEDIC* | 1 | 0.12 | 14 |
| 305 | *ADV-PRO PR S-PRO S V* | 1 | 0.12 | 14 |
| 306 | *CONJ S-PRO PART V PART CONJ S-PRO* | 1 | 0.12 | 14 |
| 307 | *CONJ S-PRO PART V PR S* | 1 | 0.12 | 14 |
| 308 | *S-PRO V V S* | 1 | 0.12 | 14 |
| 309 | *ADV V ADV-PRO ADV* | 1 | 0.12 | 14 |
| 310 | *V ADV-PRO* | 1 | 0.12 | 14 |
| 311 | *CONJ S-PRO PR S S PR A-PRO S V* | 1 | 0.12 | 14 |
| 312 | *V ADV-PRO V* | 1 | 0.12 | 14 |
| 313 | *CONJ S-PRO PR S V ADV V* | 1 | 0.12 | 14 |
| 314 | *V A-PRO S PR S CONJ S-PRO* | 1 | 0.12 | 14 |
| 315 | *CONJ S-PRO PRAEDIC* | 1 | 0.12 | 14 |
| 316 | *V NUM S* | 1 | 0.12 | 14 |
| 317 | *CONJ S-PRO PRAEDIC PART V* | 1 | 0.12 | 14 |
| 318 | *V PART CONJ S-PRO ADV PR S-PRO S V* | 1 | 0.12 | 14 |
| 319 | *CONJ S-PRO S* | 1 | 0.12 | 14 |
| 320 | *ADV-PRO S V* | 1 | 0.12 | 14 |
| 321 | *CONJ S-PRO S PART V* | 1 | 0.12 | 14 |
| 322 | *PART PR S ADV-PRO ADV A S* | 1 | 0.12 | 14 |
| 323 | *CONJ S-PRO S V* | 1 | 0.12 | 14 |
| 324 | *PART PR S S V PART A* | 1 | 0.12 | 14 |
| 325 | *CONJ S-PRO S-PRO ADV V* | 1 | 0.12 | 14 |
| 326 | *PART PR S S-PRO V NUM* | 1 | 0.12 | 14 |
| 327 | *CONJ S-PRO S-PRO PART ADV PART V* | 1 | 0.12 | 14 |
| 328 | *PART PRAEDIC* | 1 | 0.12 | 14 |
| 329 | *CONJ S-PRO S-PRO V* | 1 | 0.12 | 14 |
| 330 | *PART S* | 1 | 0.12 | 14 |
| 331 | *CONJ S-PRO S-PRO V PART S* | 1 | 0.12 | 14 |
| 332 | *PART S ADV-PRO PART PART A S PART A* | 1 | 0.12 | 14 |
| 333 | *CONJ S-PRO S-PRO V PART V* | 1 | 0.12 | 14 |
| 334 | *PART S PART S* | 1 | 0.12 | 14 |
| 335 | *CONJ S-PRO S-PRO V PR A-PRO PART S* | 1 | 0.12 | 14 |
| 336 | *PART S PR S-PRO ADV ADV V V* | 1 | 0.12 | 14 |
| 337 | *CONJ S-PRO S-PRO V S* | 1 | 0.12 | 14 |
| 338 | *ADV-PRO ADV S V* | 1 | 0.12 | 14 |
| 339 | *CONJ S-PRO S-PRO V V PART* | 1 | 0.12 | 14 |
| 340 | *PART S-PRO ADV ADV V PR S-PRO A-PRO* | 1 | 0.12 | 14 |
| 341 | *A PART V S S* | 1 | 0.12 | 14 |
| 342 | *PART S-PRO ADV S-PRO V* | 1 | 0.12 | 14 |
| 343 | *ADV V NUM S A* | 1 | 0.12 | 14 |
| 344 | *PART S-PRO ADV V PR A-PRO S* | 1 | 0.12 | 14 |
| 345 | *CONJ S-PRO V PARENTH V NUM S ADV NUM S CONJ S* | 1 | 0.12 | 14 |
| 346 | *PART S-PRO ADV-PRO A PARENTH* | 1 | 0.12 | 14 |
| 347 | *CONJ S-PRO V PART* | 1 | 0.12 | 14 |
| 348 | *PART S-PRO ADV-PRO PART A-PRO ADV V* | 1 | 0.12 | 14 |
| 349 | *ADV V S-PRO ADV V PRAEDIC PART* | 1 | 0.12 | 14 |
| 350 | *PART S-PRO CONJ V ADV-PRO* | 1 | 0.12 | 14 |
| 351 | *CONJ S-PRO V PR A-PRO S ADV CONJ V* | 1 | 0.12 | 14 |
| 352 | *PART S-PRO PART ADV-PRO S-PRO S-PRO A-PRO V* | 1 | 0.12 | 14 |
| 353 | *CONJ S-PRO V S-PRO* | 1 | 0.12 | 14 |
| 354 | *ADV-PRO ADV V S-PRO* | 1 | 0.12 | 14 |
| 355 | *CONJ S-PRO V V* | 1 | 0.12 | 14 |
| 356 | *PART S-PRO PART S-PRO V* | 1 | 0.12 | 14 |
| 357 | *A PART V S-PRO* | 1 | 0.12 | 14 |
| 358 | *PART S-PRO PART V S-PRO* | 1 | 0.12 | 14 |
| 359 | *ADV-PRO A S PART A* | 1 | 0.12 | 14 |
| 360 | *PART S-PRO PRAEDIC PR S* | 1 | 0.12 | 14 |
| 361 | *CONJ V ADV ADV-PRO PART* | 1 | 0.12 | 14 |
| 362 | *PART S-PRO S-PRO PARENTH S S* | 1 | 0.12 | 14 |
| 363 | *CONJ V ADV-PRO S A-PRO* | 1 | 0.12 | 14 |
| 364 | *ADV-PRO ADV-PRO CONJ S-PRO V* | 1 | 0.12 | 14 |
| 365 | *CONJ V CONJ S-PRO* | 1 | 0.12 | 14 |
| 366 | *PART S-PRO V A* | 1 | 0.12 | 14 |
| 367 | *PART S-PRO V PART A PR S* | 1 | 0.12 | 14 |
| 368 | *CONJ V PRAEDIC-PRO* | 1 | 0.12 | 14 |
| 369 | *PART S-PRO V S-PRO S A-PRO* | 1 | 0.12 | 14 |
| 370 | *ADV-PRO A V S* | 1 | 0.12 | 14 |
| 371 | *CONJ V S ADV* | 1 | 0.12 | 14 |
| 372 | *PART V A-PRO S* | 1 | 0.12 | 14 |
| 373 | *CONJ V S PR S* | 1 | 0.12 | 14 |
| 374 | *PART V PART INTJ* | 1 | 0.12 | 14 |
| 375 | *CONJ V S-PRO* | 1 | 0.12 | 14 |
| 376 | *ADV-PRO A-PRO V S A-PRO PART V V PART PR A-PRO S* | 1 | 0.12 | 14 |
| 377 | *PART V S-PRO PART V* | 1 | 0.12 | 14 |
| 378 | *CONJ V V* | 1 | 0.12 | 14 |
| 379 | *PART V V* | 1 | 0.12 | 14 |
| 380 | *INTJ ADV-PRO V* | 1 | 0.12 | 14 |
| 381 | *PR A S V PART A-PRO S* | 1 | 0.12 | 14 |
| 382 | *INTJ CONJ S-PRO V V PART S-PRO* | 1 | 0.12 | 14 |
| 383 | *PR A-PRO S-PRO S PART CONJ V* | 1 | 0.12 | 14 |
| 384 | *INTJ PART V* | 1 | 0.12 | 14 |
| 385 | *PR NUM V* | 1 | 0.12 | 14 |
| 386 | *INTJ PR S V S* | 1 | 0.12 | 14 |
| 387 | *PR S PART A PART S V* | 1 | 0.12 | 14 |
| 388 | *INTJ S A-PRO A* | 1 | 0.12 | 14 |
| 389 | *PR S PART V ADV* | 1 | 0.12 | 14 |
| 390 | *INTJ S CONJ S-PRO V S* | 1 | 0.12 | 14 |
| 391 | *PR S PR S-PRO ADV A S-PRO* | 1 | 0.12 | 14 |
| 392 | *INTJ S S-PRO* | 1 | 0.12 | 14 |
| 393 | *PR S S-PRO PARENTH PART S V V CONJ PR S CONJ S-PRO* | 1 | 0.12 | 14 |
| 394 | *ADV-PRO ADV A PRAEDIC*  | 1 | 0.12 | 14 |
| 395 | *PR S S-PRO V* | 1 | 0.12 | 14 |
| 396 | *INTJ S-PRO S-PRO V* | 1 | 0.12 | 14 |
| 397 | *ADV-PRO CONJ V* | 1 | 0.12 | 14 |
| 398 | *INTJ V* | 1 | 0.12 | 14 |
| 399 | *PR S-PRO ADV-PRO ADV PART A-PRO V* | 1 | 0.12 | 14 |
| 400 | *PR S-PRO ADV-PRO V* | 1 | 0.12 | 14 |
| 401 | *NUM PART V* | 1 | 0.12 | 14 |
| 402 | *PR S-PRO CONJ ADV-PRO S PART V S* | 1 | 0.12 | 14 |
| 403 | *NUM S PR S V* | 1 | 0.12 | 14 |
| 404 | *PR S-PRO PART A-PRO S V PARENTH V* | 1 | 0.12 | 14 |
| 405 | *PARENTH CONJ PR S V PART PART V* | 1 | 0.12 | 14 |
| 406 | *PR S-PRO PR S S NUM CONJ PR S* | 1 | 0.12 | 14 |
| 407 | *PARENTH CONJ S-PRO S-PRO V S* | 1 | 0.12 | 14 |
| 408 | *PR S-PRO PRAEDIC PR S A S* | 1 | 0.12 | 14 |
| 409 | *PARENTH PART PR S S-PRO V PART A-PRO S* | 1 | 0.12 | 14 |
| 410 | *PR S-PRO S ADV-PRO ADV* | 1 | 0.12 | 14 |
| 411 | *PARENTH S-PRO ADV S-PRO A* | 1 | 0.12 | 14 |
| 412 | *PR S-PRO S V* | 1 | 0.12 | 14 |
| 413 | *PARENTH S-PRO ADV S-PRO V* | 1 | 0.12 | 14 |
| 414 | *PR S-PRO V S A-PRO V CONJ* | 1 | 0.12 | 14 |
| 415 | *PARENTH S-PRO ADV V* | 1 | 0.12 | 14 |
| 416 | *PRAEDIC ADV V ADV PR S PR A-PRO S* | 1 | 0.12 | 14 |
| 417 | *PARENTH S-PRO PR S V* | 1 | 0.12 | 14 |
| 418 | *ADV-PRO PARENTH ADV S V* | 1 | 0.12 | 14 |
| 419 | *PART* | 1 | 0.12 | 14 |
| 420 | *PRAEDIC S* | 1 | 0.12 | 14 |
| 421 | *PART A S PART A-PRO V* | 1 | 0.12 | 14 |
| 422 | *PRAEDIC S-PRO PART V* | 1 | 0.12 | 14 |
| 423 | *PART ADV* | 1 | 0.12 | 14 |
| 424 | *PRAEDIC S-PRO V PR S* | 1 | 0.12 | 14 |
| 425 | *PART ADV ADV* | 1 | 0.12 | 14 |
| 426 | *PRAEDIC V ADV* | 1 | 0.12 | 14 |
| 427 | *PART ADV ADV-PRO S-PRO PART V* | 1 | 0.12 | 14 |
| 428 | *PRAEDIC V PART S-PRO* | 1 | 0.12 | 14 |
| 429 | *PART ADV S PART V* | 1 | 0.12 | 14 |
| 430 | *S ADV PART V* | 1 | 0.12 | 14 |
| 431 | *PART ADV S-PRO PR A-PRO S ADV ADV V* | 1 | 0.12 | 14 |
| 432 | *S ADV V PART* | 1 | 0.12 | 14 |
| 433 | *PART ADV-PRO A-PRO V* | 1 | 0.12 | 14 |
| 434 | *S ADV-PRO NUM* | 1 | 0.12 | 14 |
| 435 | *PART ADV-PRO PART A V S* | 1 | 0.12 | 14 |
| 436 | *S ADV-PRO S-PRO V* | 1 | 0.12 | 14 |
| 437 | *PART ADV-PRO PR A S PARENTH S V V* | 1 | 0.12 | 14 |
| 438 | *S A-PRO PRAEDIC ADV* | 1 | 0.12 | 14 |
| 439 | *PART ADV-PRO PR A S S ADV S-PRO V* | 1 | 0.12 | 14 |
| 440 | *S CONJ V ADV* | 1 | 0.12 | 14 |
| 441 | *PART ADV-PRO PR S PR A-PRO S V* | 1 | 0.12 | 14 |
| 442 | *S PART A* | 1 | 0.12 | 14 |
| 443 | *PART ADV-PRO S ADV A V* | 1 | 0.12 | 14 |
| 444 | *S PART PART PART V* | 1 | 0.12 | 14 |
| 445 | *PART ADV-PRO S PART V* | 1 | 0.12 | 14 |
| 446 | *S PART PRAEDIC PART V* | 1 | 0.12 | 14 |
| 447 | *PART ADV-PRO S-PRO A* | 1 | 0.12 | 14 |
| 448 | *S PART S-PRO CONJ V* | 1 | 0.12 | 14 |
| 449 | *PART ADV-PRO V* | 1 | 0.12 | 14 |
| 450 | *S PR A-PRO S V* | 1 | 0.12 | 14 |
| 451 | *PART ANUM* | 1 | 0.12 | 14 |
| 452 | *S PRAEDIC S* | 1 | 0.12 | 14 |
| 453 | *PART A-PRO PART V* | 1 | 0.12 | 14 |
| 454 | *S S-PRO ADV S-PRO V ADV* | 1 | 0.12 | 14 |
| 455 | *PART A-PRO S NUM V* | 1 | 0.12 | 14 |
| 456 | *S S-PRO PART V* | 1 | 0.12 | 14 |
| 457 | *PART A-PRO S S-PRO PART V* | 1 | 0.12 | 14 |
| 458 | *S S-PRO V* | 1 | 0.12 | 14 |
| 459 | *PART A-PRO S S-PRO V* | 1 | 0.12 | 14 |
| 460 | *S S-PRO V PR A S S* | 1 | 0.12 | 14 |
| 461 | *PART A-PRO S S-PRO V PR S-PRO* | 1 | 0.12 | 14 |
| 462 | *S V ADV PR S* | 1 | 0.12 | 14 |
| 463 | *PART A-PRO V A S* | 1 | 0.12 | 14 |
| 464 | *PART A-PRO V ADV S* | 1 | 0.12 | 14 |
| 465 | *S V V* | 1 | 0.12 | 14 |
| 466 | *PART CONJ ADV-PRO PART S-PRO S-PRO V* | 1 | 0.12 | 14 |
| 467 | *S-PRO A ADV-PRO V S* | 1 | 0.12 | 14 |
| 468 | *PART CONJ ADV-PRO S V PR S* | 1 | 0.12 | 14 |
| 469 | *S-PRO A S V PR S* | 1 | 0.12 | 14 |
| 470 | *PART CONJ NUM S V* | 1 | 0.12 | 14 |
| 471 | *S-PRO ADV ADV-PRO V* | 1 | 0.12 | 14 |
| 472 | *PART CONJ NUM S-PRO V* | 1 | 0.12 | 14 |
| 473 | *S-PRO ADV PART V PR A-PRO A S* | 1 | 0.12 | 14 |
| 474 | *PART CONJ PART S-PRO PART V S A-PRO PR S CONJ A-PRO PR S CONJ S* | 1 | 0.12 | 14 |
| 475 | *S-PRO ADV PR A S V* | 1 | 0.12 | 14 |
| 476 | *PART CONJ PART S-PRO S-PRO PART V* | 1 | 0.12 | 14 |
| 477 | *ADV-PRO PART S-PRO ADV PR S PART S V* | 1 | 0.12 | 14 |
| 478 | *PART CONJ PART S-PRO S-PRO V* | 1 | 0.12 | 14 |
| 479 | *S-PRO ADV V V A-PRO S* | 1 | 0.12 | 14 |
| 480 | *PART CONJ PR A-PRO S S V* | 1 | 0.12 | 14 |
| 481 | *S-PRO ADV-PRO PART ADV A* | 1 | 0.12 | 14 |
| 482 | *PART CONJ S ADV V PR S-PRO S V* | 1 | 0.12 | 14 |
| 483 | *S-PRO ADV-PRO PR S-PRO V* | 1 | 0.12 | 14 |
| 484 | *PART CONJ S A-PRO A-PRO S PART V* | 1 | 0.12 | 14 |
| 485 | *S-PRO ADV-PRO V ADV* | 1 | 0.12 | 14 |
| 486 | *PART CONJ S PART S-PRO ADV V* | 1 | 0.12 | 14 |
| 487 | *S-PRO A-PRO S V* | 1 | 0.12 | 14 |
| 488 | *PART CONJ S S PR S V* | 1 | 0.12 | 14 |
| 489 | *S-PRO CONJ NUM NUM V* | 1 | 0.12 | 14 |
| 490 | *PART CONJ S-PRO A-PRO V* | 1 | 0.12 | 14 |
| 491 | *S-PRO CONJ PR S V* | 1 | 0.12 | 14 |
| 492 | *PART CONJ S-PRO PART ADV ADV-PRO V* | 1 | 0.12 | 14 |
| 493 | *S-PRO CONJ V* | 1 | 0.12 | 14 |
| 494 | *PART CONJ S-PRO PART V PR S-PRO S* | 1 | 0.12 | 14 |
| 495 | *S-PRO NUM CONJ S PRAEDIC V* | 1 | 0.12 | 14 |
| 496 | *PART CONJ S-PRO PR S ADV V PR A-PRO S* | 1 | 0.12 | 14 |
| 497 | *S-PRO PART ADV A-PRO S* | 1 | 0.12 | 14 |
| 498 | *PART CONJ S-PRO S-PRO S-PRO PART V* | 1 | 0.12 | 14 |
| 499 | *S-PRO PART ADV-PRO V* | 1 | 0.12 | 14 |
| 500 | *PART CONJ S-PRO V* | 1 | 0.12 | 14 |
| 501 | *S-PRO PART A-PRO S S* | 1 | 0.12 | 14 |
| 502 | *PART CONJ S-PRO V PART* | 1 | 0.12 | 14 |
| 503 | *S-PRO PART NUM S ADV-PRO PART V V* | 1 | 0.12 | 14 |
| 504 | *PART CONJ S-PRO V PR S S* | 1 | 0.12 | 14 |
| 505 | *ADV PR NUM NUM PR S V S PR S V CONJ PR S-PRO ADV-PRO V* | 1 | 0.12 | 14 |
| 506 | *PART CONJ V S A-PRO A* | 1 | 0.12 | 14 |
| 507 | *S-PRO PART PR A S V S* | 1 | 0.12 | 14 |
| 508 | *PART ADV-PRO S-PRO S-PRO V ADV-PRO* | 1 | 0.12 | 14 |
| 509 | *S-PRO PART PR S-PRO PR S V* | 1 | 0.12 | 14 |
| 510 | *PART NUM NUM V* | 1 | 0.12 | 14 |
| 511 | *S-PRO PART PRAEDIC ADV-PRO* | 1 | 0.12 | 14 |
| 512 | *PART PARENTH* | 1 | 0.12 | 14 |
| 513 | *ADV ADV V A-PRO PR S* | 1 | 0.12 | 14 |
| 514 | *PART PARENTH ADV-PRO S-PRO ADV-PRO V PR A-PRO S* | 1 | 0.12 | 14 |
| 515 | *S-PRO PART V PR A-PRO CONJ A-PRO* | 1 | 0.12 | 14 |
| 516 | *PART PARENTH CONJ S-PRO ADV PART V* | 1 | 0.12 | 14 |
| 517 | *S-PRO PR A-PRO S PART V* | 1 | 0.12 | 14 |
| 518 | *PART PARENTH PRAEDIC CONJ ADV-PRO V S-PRO ADV* | 1 | 0.12 | 14 |
| 519 | *S-PRO PR S ADV V S* | 1 | 0.12 | 14 |
| 520 | *PART PARENTH S-PRO CONJ PART V* | 1 | 0.12 | 14 |
| 521 | *ADV-PRO PR A-PRO S PRAEDIC* | 1 | 0.12 | 14 |
| 522 | *PART PARENTH S-PRO PARENTH CONJ PART PRAEDIC* | 1 | 0.12 | 14 |
| 523 | *PART PARENTH V S* | 1 | 0.12 | 14 |
| 524 | *S-PRO PR S-PRO ADV V ADV-PRO* | 1 | 0.12 | 14 |
| 525 | *PART PARENTH V V* | 1 | 0.12 | 14 |
| 526 | *S-PRO PR S-PRO CONJ S-PRO PR S-PRO V* | 1 | 0.12 | 14 |
| 527 | *PART PART ADV PART PARENTH PRAEDIC* | 1 | 0.12 | 14 |
| 528 | *S-PRO S* | 1 | 0.12 | 14 |
| 529 | *PART PART ADV PRAEDIC S-PRO V* | 1 | 0.12 | 14 |
| 530 | *S-PRO S PART V* | 1 | 0.12 | 14 |
| 531 | *PART PART ADV-PRO S S S-PRO V ADV PR S A S CONJ PART* | 1 | 0.12 | 14 |
| 532 | *S-PRO S-PRO ADV* | 1 | 0.12 | 14 |
| 533 | *PART PART A-PRO V PR S S* | 1 | 0.12 | 14 |
| 534 | *S-PRO S-PRO A-PRO S PRAEDIC* | 1 | 0.12 | 14 |
| 535 | *V PR S PR S* | 1 | 0.12 | 14 |
| 536 | *S-PRO S-PRO PART V PARENTH* | 1 | 0.12 | 14 |
| 537 | *V PR S S-PRO PART A-PRO* | 1 | 0.12 | 14 |
| 538 | *S-PRO S-PRO PART V PR S-PRO V INTJ* | 1 | 0.12 | 14 |
| 539 | *V PR S-PRO S* | 1 | 0.12 | 14 |
| 540 | *ADV-PRO PR S S V S* | 1 | 0.12 | 14 |
| 541 | *ADV ADV-PRO PART PRAEDIC V CONJ PART* | 1 | 0.12 | 14 |
| 542 | *V S PARENTH* | 1 | 0.12 | 14 |
| 543 | *ADV-PRO PR S V S* | 1 | 0.12 | 14 |
| 544 | *ADV-PRO ADV ANUM V* | 1 | 0.12 | 14 |
| 545 | *V S-PRO V ADV* | 1 | 0.12 | 14 |
| 546 | *S-PRO V PART V PART* | 1 | 0.12 | 14 |
| 547 | *ADV-PRO ADV PR S-PRO V PART INTJ* | 1 | 0.12 | 14 |
| 548 | *PART PART S-PRO PART V* | 1 | 0.12 | 14 |
| 549 | *S-PRO V PR S-PRO S S S* | 1 | 0.12 | 14 |
| 550 | *PART PART S-PRO PART V ADV-PRO* | 1 | 0.12 | 14 |
| 551 | *S-PRO V S-PRO CONJ PRAEDIC* | 1 | 0.12 | 14 |
| 552 | *PART PART S-PRO S-PRO CONJ PART V ADV A* | 1 | 0.12 | 14 |
| 553 | *S-PRO V V CONJ* | 1 | 0.12 | 14 |
| 554 | *PART PART S-PRO S-PRO V S* | 1 | 0.12 | 14 |
| 555 | *ADV A-PRO PART A-PRO A S A PART S* | 1 | 0.12 | 14 |
| 556 | *PART PART S-PRO V* | 1 | 0.12 | 14 |
| 557 | *PART PART S-PRO V PR S-PRO* | 1 | 0.12 | 14 |
| 558 | *V ADV-PRO A-PRO A S* | 1 | 0.12 | 14 |
| 559 | *PART PART S-PRO V S-PRO ADV ADV V* | 1 | 0.12 | 14 |
| 560 | *V ADV-PRO S-PRO* | 1 | 0.12 | 14 |
| 561 | *A S* | 1 | 0.12 | 14 |
| 562 | *V ADV-PRO V S PR S* | 1 | 0.12 | 14 |
| 563 | *PART PART V A-PRO S PART PART S* | 1 | 0.12 | 14 |
| 564 | *V A-PRO S* | 1 | 0.12 | 14 |
| 565 | *PART PART V PART* | 1 | 0.12 | 14 |
| 566 | *V CONJ V* | 1 | 0.12 | 14 |
| 567 | *PART PART V PR S S* | 1 | 0.12 | 14 |
| 568 | *ADV-PRO S PART* | 1 | 0.12 | 14 |
| 569 | *PART PART V S* | 1 | 0.12 | 14 |
| 570 | *ADV-PRO S PART PART V PART* | 1 | 0.12 | 14 |
| 571 | *PART PART V V S INTJ* | 1 | 0.12 | 14 |
| 572 | *V PART ADV-PRO* | 1 | 0.12 | 14 |
| 573 | *PART PR A S S-PRO V* | 1 | 0.12 | 14 |
| 574 | *V PART V PR S ADV* | 1 | 0.12 | 14 |
| 575 | *PART PR A-PRO S CONJ V ADV CONJ PR NUM PART V* | 1 | 0.12 | 14 |
| 576 | *V PR NUM S* | 1 | 0.12 | 14 |
| 577 | *PART PR A-PRO S PRAEDIC* | 1 | 0.12 | 14 |
| 578 | *V PR S ADV* | 1 | 0.12 | 14 |
| 579 | *PART PR A-PRO V S PART* | 1 | 0.12 | 14 |
| 580 | *PART PR S ADV PR S PART PRAEDIC* | 1 | 0.12 | 14 |
| 581 | *V PR S V V* | 1 | 0.12 | 14 |
| 582 | *V V INTJ* | 1 | 0.12 | 14 |
| 583 | *ADV CONJ V PR A-PRO S PR A S* | 1 | 0.12 | 14 |
| 584 | *A PARENTH* | 1 | 0.12 | 14 |
| 585 | *PART PART PART S-PRO CONJ V* | 1 | 0.12 | 14 |
| 586 | *V S S V* | 1 | 0.12 | 14 |
| 587 | *PART PART PR A-PRO S S V ADV ADV-PRO* | 1 | 0.12 | 14 |
| 588 | *V S-PRO PR S* | 1 | 0.12 | 14 |
| 589 | *PART PART S-PRO* | 1 | 0.12 | 14 |
| 590 | *PART PART S-PRO A* | 1 | 0.12 | 14 |
| 591 | *PART PART S-PRO ADV V V* | 1 | 0.12 | 14 |
| 592 | *PART PART CONJ A-PRO V* | 1 | 0.12 | 14 |
| 593 | *V V PART PR S-PRO A-PRO S* | 1 | 0.12 | 14 |
| 594 | *PART PART CONJ PR S-PRO S-PRO A-PRO S V* | 1 | 0.12 | 14 |
| 595 | *S-PRO ADV V PR S-PRO* | 1 | 0.12 | 14 |
| 596 | *S-PRO A S V* | 1 | 0.12 | 14 |
| 597 | *ADV-PRO S-PRO V PART* | 1 | 0.12 | 14 |
| 598 | *PART PART CONJ S-PRO PARENTH S-PRO V ADV* | 1 | 0.12 | 14 |

1. URL: http://www.ruscorpora.ru/index.html. [↑](#footnote-ref-1)
2. URL: http://www.lingexp.uni-tuebingen.de/sfb441/b1/rus/korpora.html. [↑](#footnote-ref-2)
3. URL: http://www.helsinki.fi/venaja/russian/e-material/hanco/index.htm. [↑](#footnote-ref-3)
4. URL: http://www.philol.msu.ru/~lex/corpus/. [↑](#footnote-ref-4)
5. URL: http://www.narusco.ru/. [↑](#footnote-ref-5)
6. URL: http://www.webcorpora.ru/. [↑](#footnote-ref-6)
7. URL: https://www.sketchengine.co.uk/rutenten-corpus. [↑](#footnote-ref-7)
8. URL: http://opencorpora.org/. [↑](#footnote-ref-8)
9. URL: http://web-corpora.net/learner\_corpus. [↑](#footnote-ref-9)
10. URL: http://spokencorpora.ru/. [↑](#footnote-ref-10)