

УДК 81'27

ЧЕЛОВЕК И ТЕКСТ: АНТРОПОЦЕНТРИЧЕСКИЙ ПОДХОД К ИССЛЕДОВАНИЮ¹

Елена Валентиновна Ерофеева

профессор кафедры общего и славянского языкознания

Пермский государственный университет

614990, Пермь, ул. Букирева, 15. erofeeva@newmail.ru

Тамара Ивановна Ерофеева

профессор кафедры общего и славянского языкознания

Пермский государственный университет

614990, Пермь, ул. Букирева, 15. erofkaf@rambler.ru

В статье рассматриваются результаты исследования влияния социальных факторов на речевую продукцию говорящих на синтаксическом уровне. На материале спонтанной монологической речи исследуются такие синтаксические параметры, как «длина предложения» и «синтаксическая сложность». Для сопоставления привлекаются данные супraseгментного анализа спонтанной речи, которые непосредственно связаны с синтаксическим уровнем, а именно: средняя длина фразы и средняя длина синтагмы в текстах. Результаты анализа данных по трем выборкам испытуемых убедительно свидетельствуют о том, что на синтаксический уровень оказывают существенное влияние факторы «образование» и «гендер», при этом градации фактора «образование» существенно влияют и на длину предложения, и на синтаксическую сложность, а градации фактора «гендер» – только на длину предложения.

Ключевые слова: спонтанная речь; социальные факторы; синтаксис; предложение; синтагма; фраза; статистическое моделирование.

Данная статья созвучна идеям, отраженным в книге Л.Н.Мурзина и А.С.Штерн «Текст и его восприятие» [Мурзин, Штерн 1991].

Вслед за Л.Н.Мурзиным текст понимается и как вертикальное и горизонтальное соединение фонем, морфем, слов, предложений и других единиц, и как факт культуры. Неслучайно затем, выдвигая тезис о двух функциях языка: 1) сообщении о чем-то объективном; 2) информации не о мире вообще, а о человеке [Фатическое поле языка 1998: 11], Л.Н.Мурзин заключает: «Без второй функции язык немислим» [там же: 12].

Данный тезис актуален и для современной лингвистики, прежде всего для тех ее направлений, которые связаны с исследованием функционирования языка. Сегодня «задачей когнитивистики текста признается исследование структуры знаний индивида, особое внимание к механизмам порождения и понимания высказываний, допущение того, что названные процессы, протекающие в рамках структуры процесса текущей мыслительной деятельности, включают языко-

вые и неязыковые элементы» [Кинчина, Мусихина 2003; см. также Виноград 1983].

Социолингвистический подход, будучи одной из ветвей лингвистики, ориентированных на изучение речевого материала, позволяет интерпретировать текст как произведение конкретного говорящего, обладающего определенными социальными характеристиками.

Давно известно, что для каждого говорящего выбор языка, стиля, конкретной грамматики или лексической единицы предсказуем на основе ситуации, жанра речи или таких признаков, как пол или положение собеседника [Эрвин-Трипп 1975].

Несомненный интерес вызывает вопрос о влиянии гендера на построение текста. Однако значительное количество работ в области гендерной лингвистики, рассматривающих социальную природу языка мужчин и женщин, не позволяет однозначно понять это явление. Это связано прежде всего с разнородностью анализируемого материала, а также с тем, что фактор «гендер» неоднороден и сопряжен с такими факторами,

как речевая ситуация, тематика разговора, возраст и образование.

В данном исследовании рассматривались факторы возраст, место рождения, образование, гендер. Материалом исследования послужили устные спонтанные монологические тексты.

Исследование проводилось в два этапа. На первом этапе рассматривались монологи на тему «Комическая жизненная ситуация» и изучалась только гендерная специфика речи информантов. В данном случае было рассмотрено 12 текстов информантов-женщин и 12 – информантов-мужчин. Все информанты – студенты Пермского государственного университета, преимущественно гуманитарных факультетов, образование неоконченное высшее, возрастная группа одинаковая (18–22 года).

В полученных монологах изучались синтаксические признаки. В качестве результирующих исследовались следующие признаки:

- длина предложения в словах;
- синтаксическая сложность предложения.

Подобного рода исследования спонтанной речи всегда сопряжены с определенными методическими сложностями, связанными с невозможностью простого «перенесения» грамматических категорий, разработанных для языка, на стихию речи. В данном случае трудность заключалась прежде всего в нахождении границ предложения. Под простым предложением мы понимали минимальную предикативную единицу [Белошапкина 1977; Потапова 1972; Филиппов 1993], а под сложным – синтаксическую единицу, «компонентами которой являются предложения, связанные между собой синтаксической связью» [Современный русский язык 1981: 366].

Поскольку проблема определения слова также сложна, в целях упрощения методики отдельным словом считалась единица письменной речи от пробела до пробела [Ерофеева Т.И. 2004: 156]. Такая методика используется во многих лингвистических исследованиях [см., например: Фонетика спонтанной речи 1988: 220]. Средняя длина предложения в тексте определялась делением суммы всех слов на количество предложений.

При определении синтаксической сложности учитывалось количество сочинительных и подчинительных связей внутри сложного предложения. Каждому типу синтаксической конструкции приписывался отдельный коэффициент сложности: простое предложение – 1; сложносочиненное – 1,5; сложноподчиненное – 1,5; бессоюзное сложное – 2; сложноподчиненное с последовательным подчинением – 2,5; сложносочиненное, объединенное со сложноподчиненным – 3 [Еро-

феева Т.И. 2004: 156]. Средняя синтаксическая сложность конструкций текста (\bar{k}) рассчитывалась как средняя взвешенная:

$$\bar{k} = \frac{\sum k_i \cdot m_i}{\sum m_i},$$

где k_i – коэффициент сложности определенной конструкции, m_i – количество соответствующих конструкций в тексте, $\sum m_i$ – общее количество предложений-конструкций в тексте.

Данные о средней длине предложения и средней сложности синтаксической конструкции, полученные на первом этапе исследования, отражены в таблице 1 (информанты с одинаковым уровнем образования и возрастом имеют одинаковые номера).

Таблица 1
Средняя длина предложения (l) и показатель синтаксической сложности (k) текстов мужчин и женщин

№ информанта	Женщины		Мужчины	
	l	k	l	k
1	11,7	1,2	9	1,2
2	14,2	1,5	7,2	1,2
3	11,5	0,5	10,3	1,3
4	10	1,2	6,4	1
5	11	1,3	15,3	1,4
6	8,8	1,3	16	1,4
7	19,2	1,8	19,6	2
8	31,7	1,8	17,3	2
9	37,5	1,9	16,5	1,6
10	2,6	1,6	14,2	1,6
11	9,6	1,2	13,2	1,6
12	13,3	1,3	10,6	1,1
В среднем	15,1	1,4	13,0	1,5

Таблица показывает, что средняя длина предложения и средний показатель синтаксической сложности текстов мужчин и женщин имеют небольшие различия.

Представим данные на графиках (рис. 1, 2).

Как видим, в группе мужчин обнаруживается относительная однородность величины средней длины предложения, показатель которой находится в пределах от 6 до 16 слов в предложении. Он отмечен в текстах 8 из 12 мужчин информантов; в речи трех информантов (№ 10, 11, 12) длина предложения не превышала 10 слов. И только один информант (№ 7) продемонстрировал среднюю длину предложения, близкую к 20 словам – 19,6.

В группе информантов-женщин длина предложения варьирует от 2,6 до 37,5 слов и характеризуется наличием артефактов. Показатели длины предложения женщин можно разделить на три группы. Самая большая группа включает показатели 6 информантов (№ 1, 2, 3, 4, 5, 11) с длиной предложения от 10 до 14 слов. Информанты № 7, 8, 9 имеют длину предложения от 19

до 38 слов. Такую особенность синтаксической организации речи женщин можно толковать как проявления условно называемой «лингвистической несдержанности». И три информанта (№ 6, 10, 11) продемонстрировали среднюю длину предложения в монологах меньше 10 слов (2,6; 8,8; 9,6).

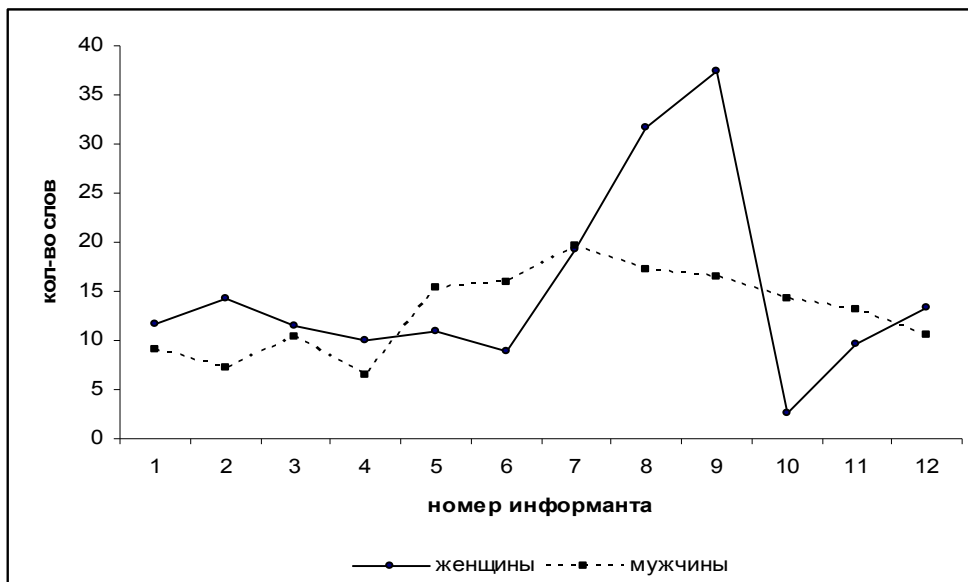


Рис. 1. Сравнение средней длины предложения мужчин и женщин при одинаковых показателях возраста и образования

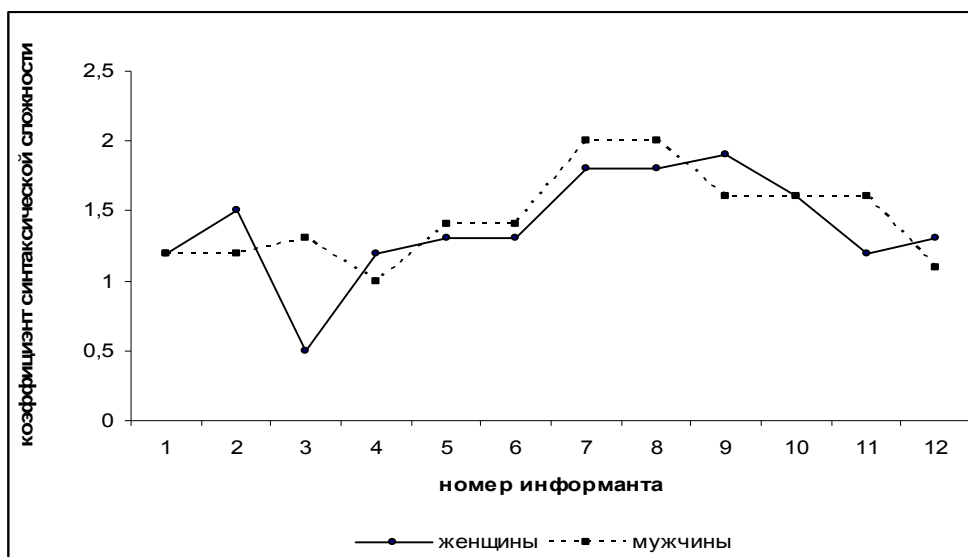


Рис. 2. Сравнение среднего показателя синтаксической сложности текстов мужчин и женщин при одинаковых показателях возраста и образования

Рассмотрим график на рис. 2, где изображен средний показатель синтаксической сложности текстов мужчин и женщин при одинаковых показателях возраста и образования. Средний показатель синтаксической сложности текстов двенадцати мужчин находится в пределах от 1 до 2.

Можно выделить две группы. 10 информантов из 12 имеют средний показатель синтаксической сложности текста в пределах от 1 до 1,6. Информанты № 7 и 8 объединяются в группу со средним показателем синтаксической сложности текстов, равным 2. Как видим, синтаксическая

Таблица 2

Длина предложения (*l*) и синтаксическая сложность (*k*) в зависимости от социальных факторов

сложность текстов мужчин достаточно однородна и выше, чем у женщин. В группе женщин средний показатель синтаксической сложности текста варьирует в большем диапазоне, чем у мужчин – от 0,5 до 1,9. Выделяются три группы. Первая (7 информантов – № 1, 2, 4, 5, 6, 11, 12) имеет показатель от 1,2 до 1,5; вторая (информанты № 7, 8, 9, 10) – от 1,6 до 1,9. И один информант 3 продемонстрировал средний показатель синтаксической сложности текста, равный 0,5.

Итак, анализ данных показывает, что при равных показателях возраста и образования наблюдаются некоторые различия в синтаксической структуре речи мужчин и женщин. У мужчин вариация количественных показателей ниже, чем у женщин: явных значительных завышений или занижений показателя длины предложения на фоне среднего не выявлено. Хотя для спонтанной речи мужчин характерна несколько бóльшая синтаксическая сложность конструкций, чем для речи женщин, эти различия оказываются настолько малы, что говорить об устойчивой разнице преждевременно. Для этого необходимо статистический анализ данных.

Такой анализ осуществлялся на втором этапе исследования, где вместе в фактором «гендер» рассматривались и другие социальные факторы говорящих: «образование», «возраст» и «место рождения».

В данном случае рассматривалась спонтанная речь (пересказ телесериала) 32 информантов. Выборка информантов была сбалансирована по 4 факторам: образованию (высшее, среднее), месту рождения (г. Пермь, районный центр/село), возрасту (группы 20–29 лет, 30–39 лет, 40–49 лет, 50 лет и старше), гендеру. Общая протяженность звучания речевых отрезков составила 160 минут. Исследовались те же два результирующих признака, что и в первой части эксперимента: средняя длина предложения и средняя синтаксическая сложность, – рассчитанные по той же методике (см. табл. 2).

Как видно из таблицы, средняя длина предложения оказывается больше в речи уроженцев краевого центра, по сравнению с речью лиц, рожденных в районных центрах и селах. Этот же параметр выше в речи мужчин, чем в речи женщин; в речи людей с высшим образованием, чем в речи людей со средним образованием. Если рассматривать возрастные группы, то наиболее длинные предложения отмечаются в речи людей 40–49 лет, а самые короткие – в старшей возрастной группе.

Факторы		Синтаксические параметры	
		<i>l</i>	<i>k</i>
Место рождения	Пермь	27,8	1,8
	Край	24,7	1,7
Гендер	Муж.	26,9	1,7
	Жен.	21,8	1,5
Образование	Высш.	28,3	1,7
	Сред.	21,0	1,4
Возраст	20-29	22,6	1,6
	30-39	23,9	1,6
	40-49	26,1	1,7
	≥50	19,8	1,5

Анализ данных по параметру «синтаксическая сложность» показывает, что наибольшая количественная разница достигается для градаций фактора «образование»: коэффициент синтаксической сложности речи лиц с высшим образованием на 0,3 выше, чем речи лиц со средним образованием. В остальных случаях разница по градациям факторов не выше 0,2.

В отличие от первого этапа, в данном случае при анализе речевой продукции информантов был применен дисперсионный анализ силы влияний (ДА), который позволяет оценить значимость влияния факторов и ранжировать их силу влияния [Штерн 1992; см. также: Ерофеева Т.И. 2004, Ерофеева Е.В. 2005].

В таблице 3 представлены ранги факторов по степени их влияния на проявление исследуемых признаков: среднюю длину предложения и синтаксическую сложность в спонтанной речи говорящих. В скобках в таблице даны несущественные факторы.

Таблица 3

Ранги факторов по силе влияния на длину предложения (*l*) и синтаксическую сложность (*k*)

Факторы	Синтаксические параметры	
	<i>l</i>	<i>k</i>
Место рождения	(3)	(4)
Гендер	2	(3)
Образование	1	1
Возраст	(4)	(2)

ДА показывает, что для признака «длина предложения» ведущим оказывается фактор «об-

разование», а на втором месте оказывается фактор «гендер». Фактор «возраст» оказывается незначимым и занимает по своему рангу последнее место. Для признака «синтаксическая сложность» существенным оказался только один из факторов – «образование». Влияние остальных факторов несущественно, хотя фактор «возраст» имеет высокий показатель силы влияния ($\eta_x^2 = 9,6$).

Таким образом, подтверждаются на другом материале предварительные данные первого этапа исследования о влиянии фактора «гендер» на среднюю длину предложения и о независимости от этого фактора синтаксической сложности высказывания.

Синтаксические характеристики спонтанной речи соотносятся с ее фонетической организацией, прежде всего фраз и синтагм (еще Л.В.Щерба называл синтагму простейшим синтаксическим целым). Поэтому данные о синтаксических характеристиках были соотнесены нами с данными о средней длине синтагм и фраз в спонтанной речи информантов с подобными социальными характеристиками.

В данном случае для анализа был взят материал хрестоматии «Городские социолекты: пермская городская речь» [Ерофеева Т.И., Ерофеева Е.В., Грачева, 2000], в которой представлены спонтанные монологи 24 информантов на тему «Работа». Выборка информантов сбалансирована по градациям факторов «возраст» (25–34 года, 35–44 года, 45–55 лет), «место рождения» (г. Пермь, города и поселки Пермского края), «образование» (среднее, высшее) и «гендер».

Рассматривались следующие показатели:

- длина фразы в синтагмах,
- длина фразы в словах,
- длина синтагмы в словах.

Полученные данные представлены в таблице 4.

Анализ данных показывает, что в спонтанной речи уроженцев Перми фразы и синтагмы оказываются более длинными, чем в речи уроженцев области, что соответствует поведению параметра «длина предложения» (см. табл. 2).

У женщин длина синтагмы в словах, а также длина фразы в словах также оказываются выше, чем у мужчин, что также согласуется с приведенными ранее данными. Длина фразы в синтагмах, напротив, больше у мужчин, что может косвенно свидетельствовать о большей синтаксической сложности фразы.

Наибольшие различия по градациям для параметров «длина фразы в словах» и «длина фразы в синтагмах» наблюдаются у фактора «обра-

зование»: люди с высшим образованием строят значительно более длинные фразы, чем люди со средним образованием. Однако длина синтагмы в словах практически не зависит от данного фактора.

Таблица 4

Длина фраз и синтагм спонтанных монологов в зависимости от социальных факторов

Факторы		Длина фразы		Длина синтагмы в словах
		в синтагмах	в словах	
Место рождения	Пермь	3,3	9,2	2,8
	Край	3,1	7,8	2,4
Гендер	Муж.	3,4	7,6	2,4
	Жен.	3,1	8,6	2,8
Образование	Высш.	3,7	10,1	2,7
	Сред.	2,8	7,2	2,5
Возраст	25-34	2,8	7,4	2,5
	35-44	3,0	8,8	2,9
	45-55	3,9	10,4	2,4

Что касается фактора «возраст», то наиболее длинные фразы так же, как и в предыдущем случае, зафиксированы в речи возрастной группы около 50 лет.

Таким образом, мы видим, что, действительно, синтаксические и фонетические супrasegmentные характеристики речи оказываются тесно связанными. Длина предложения соотносится с такой характеристикой собственно устной речи, как длина фразы в словах, а синтаксическая сложность соотносится с длиной фраз в синтагмах. При этом показательно, что результаты, полученные на разном материале и с разными группами информантов оказываются очень устойчивыми, что говорит о надежности полученных выводов.

* * *

Итак, на синтаксическом уровне наиболее важным фактором оказался фактор «образование»: у лиц с высшим образованием длина предложения и фразы, а также степень сложности выше, чем у лиц со средним образованием, что представляется закономерным. На втором месте находится фактор «гендер», который оказывает существенное влияние только на длину предложения: предложения у мужчин длиннее, чем у женщин.

Таким образом, применение статистических методов в социолингвистическом исследовании позволяет описать социальный опыт, отраженные в социальных характеристиках говорящих, в

употреблении языковых параметров. Надежность данного метода подтверждается повторяемостью результатов.

¹ Исследование выполнено при поддержке гранта РГНФ № 08-04-0004а

Список литературы

Белошапкова В.А. Современный русский язык. Синтаксис. М., 1977. 248 с.

Виноград Т. К процессуальному пониманию семантики // Новое в зарубежной лингвистике. М., 1983. Вып. 12. С. 123–170.

Ерофеева Е.В. Вероятностные структуры идиомов: социолингвистический аспект. Пермь: Изд-во Перм. ун-та. 2005. 320 с.

Ерофеева Т.И. Современная городская речь / ПГУ, ПСИ, ПССГК. Пермь, 2004. 316 с.

Ерофеева Т.И., Ерофеева Е.В., Грачева И.И. Городские социолекты: Пермская городская речь: звучащая хрестоматия. Пермь; Бохум, 2000.

Кинчина Е.В., Мусихина О.И. Тексты с позиции когнитивной лингвистики // Языковые и культурные контакты различных народов: сб. статей. Пенза, 2003. С. 118–122.

Мурзин Л.Н., Штерн А.С. Текст и его восприятие / УРГУ. Свердловск, 1991. 172 с.

Потанова Н.П. Структура предложений с глаголами, обозначающими линейное перемещение (на материале говоров северной части Пермской области) // Уч. зап. Перм. ун-та: Вопросы грамматики, 1972. № 243. С. 189–204.

Современный русский язык / под ред. В.А.Белошапковой. М., 1981. 560 с.

Фатическое поле языка (памяти профессора Л.Н.Мурзина): Межвузовский сборник научных трудов / Перм. ун-т. Пермь, 1998. 225 с.

Филиппов К.А. Проблемы синтаксиса и интонации спонтанной речи (на материале немецкого языка). СПб.: Изд-во СПб. ун-та, 1993, 164 с.

Фонетика спонтанной речи / под ред. Н.Д.Светозаровой. Л.: Изд-во Ленингр. ун-та, 1988. 248 с.

Штерн А.С. Перцептивный аспект речевой деятельности. СПб.: Изд-во СПб. ун-та, 1992. 236 с.

Эрвин-Трипп С. Социолингвистика в США // Социально-лингвистические исследования. М., 1975. С 181–199.

THE PERSON AND THE TEXT: THE ANTHROPOCENTRIC APPROACH TO RESEARCH

Elena V. Erofeeva

**Professor of General Linguistics and Slavonic Languages Department
Perm State University**

Tamara I. Erofeeva

**Professor of General Linguistics and Slavonic Languages Department
Perm State University**

The results of research of social factors influence on the syntactic level of speech production are considered in article. On material of spontaneous monologic speech such syntactic parametres as «length of the sentence» and «syntactic complexity» are investigated. The data of the suprasegmental analysis of spontaneous speech which are directly connected with syntactic level is involved for comparison. Average length of phrase and average length of syntagma in texts is considered. The results of the analysis of the data in three samples of examinees argue that factors influencing essentially the syntactic level are «education» and «gender». Gradations of the factor «education» essentially influence both length of sentence, and syntactic complexity, and gradations of the factor «gender» influence only the length of sentence.

Key words: spontaneous speech; social factors; syntax, sentence; syntagma; phrase; statistical modelling.