

УДК 81.13

DOI: 10.18101/1994-0866-2018-2-2-33-41

ЧАСТОТНЫЕ СПИСКИ КАК ИНСТРУМЕНТ ПОИСКА ЛИНГВИСТИЧЕСКИХ КОРРЕЛЯТОВ СОЦИАЛЬНЫХ РОЛЕЙ ГОВОРЯЩЕГО*

© *Попова Татьяна Ивановна*

магистрант,

Санкт-Петербургский государственный университет

Россия, 199034, Санкт-Петербург, Университетская набережная, 7/9

E-mail: tipopova13@gmail.com

В статье представлен анализ частотных списков, составленных на материале расшифровок записей одного информанта из Звукового корпуса русского языка «Один речевой день». Конкретным материалом послужили 5 эпизодов, в которых информант выступает в социальных ролях, противопоставленных друг другу по принципу асимметричности социальных отношений: роль *мужа* во время общения с женой и роль при общении с клиентами, обозначенная в корпусе как *служба сервиса*. Предметом исследования стали наиболее употребительные единицы из верхней зоны частотного списка, которые являются специфичными и тем самым могут быть интерпретированы как лингвистические корреляты социальной роли говорящего, создающие особый субкод, описанный в лингвистике только в теоретическом аспекте. При этом маркированной оказывается ситуация формального общения. Дальнейший анализ извлеченных единиц позволит определить, каково значение социальной роли и каков вес коммуникативной ситуации при переключении речевого поведения.

Ключевые слова: русская разговорная речь; корпус; социолингвистика; социальная роль; коммуникативная ситуация; лингвистические корреляты; частотные списки; дискурсивный анализ; теория вежливости; коммуникативные стратегии.

Задача каждого участника коммуникации – оказать воздействие на слушающего, зачастую только лишь донесением своей мысли, передачей информации. Способ решения этой задачи зависит от коммуникативной ситуации, которая всегда конкретна и имеет динамичный характер, но, с другой стороны, она всегда представляет собой вариант уже существующего типа ситуации.

Сам процесс усвоения языка тесно связан с усвоением правил (не только грамматических) его использования. «Способность сделать правильный выбор, осуществить необходимое ролевое речевое переключение, адекватно заявив о себе в различных ситуациях общения, является показателем сформированной коммуникативной личности, <...> позволяющей человеку быть не просто говорящим, а членом социально обусловленной системы общения» [9]. Говорящий усваивает абстрактные образцы речевого поведения, реализация которых напрямую связана с понятием социальной роли говорящего.

Под социальной ролью (СР) будем понимать «нормативно одобренный обществом образ поведения, ожидаемый от каждого, занимающего данную социальную позицию» [4, с. 23]. Образцовый характер и стереотипность роли

* Исследование выполнено при поддержке гранта РФФИ № 17-29-09175 «Диагностические признаки социолингвистической вариативности повседневной русской речи (на материале звукового корпуса)».

подтверждается тем, что она определяется с помощью уже существующих в языке обозначений: *муж, сын, клиент, пациент* и другие [5, с. 427].

Смена ролей изменяет ситуацию, отражаясь на выборе языковых средств [10]. Переключение с одних стереотипов на другие касается и речевого поведения: «...ситуативная вариативность проявляется в использовании тех или иных языковых средств – отдельных единиц или целых систем/подсистем в зависимости от социальной ситуации» [10, с. 79], но есть мнение, что речь в таком случае идет только о «стандартных ролевых *инвариантах* (курсив мой. – Т. П.) и различных исполнительских вариантах, в каждом из которых личность предстает в новом образе» [9].

Набор подобного рода единиц называют социолингвистическими (ситуативными) переменными, в данном исследовании в этом значении будет использоваться термин *лингвистические корреляты*. Классификации этих переменных в литературе обнаружить не удалось, хотя исследователи отмечают, что «для социолингвистики релевантны не любые роли, а лишь те из них, которые имеют эксплицитно выраженные языковые или речевые корреляты» [10, с. 82]. Стоит отметить, что в работах на эту тему встречаются мысли о том, что специфика речи задается не ролью: «дело, как кажется, не просто в характере социальной роли, а в целевой установке при ее исполнении в той или иной коммуникативной ситуации» [5, с. 435].

Цель настоящего исследования – составление и анализ частотных списков (ЧС), в верхней зоне которых могут оказаться единицы, маркирующие речевое переключение при смене стереотипов поведения в разных социальных ролях одного и того же говорящего.

Для достижения поставленной цели используются данные Звукового корпуса русского языка «Один речевой день» (ОРД), который является уникальным источником материала для анализа спонтанной разговорной русской речи. На основании анкет, которые заполняют участники записи, были составлены базы данных, где содержатся характеристики информантов (пол, возраст, уровень образования, место рождения и др.), а также данные о тех эпизодах, из которых состоит их «речевой день» (место, время, коммуниканты, тема и др.). Этот материал позволил сформировать классификацию социальных ролей говорящего в ОРД в зависимости от коммуникативной ситуации (см. рис. 1).

Около трети данных (29,4%) приходится на роль *коллег*, но отношения между коллегами оказываются неоднозначны: нередко вместе работают близкие друзья, которые могут находиться в отношениях «начальник-подчиненный», что не позволяет четко интерпретировать полученные данные. Зато ситуации общения, в которых говорящий выступает в социальной роли, например, *мужа* (МУ – 10%) и *представителя сервисной службы* (СС – 14%), оказываются противопоставлены друг другу по принципу формальности/неформальности или асимметричности/симметричности ролевых отношений, более того, процентное соотношение в корпусе различается для выбранных ролей незначительно [1, с. 201–202]. Такое положение дел позволяет предположить возможность выявления лингвистических коррелятов для двух ролей, близких по значимости в повседневном общении.

<p>Родственники:</p> <ul style="list-style-type: none"> • МУ – муж, • ЖЕ – жена • ОТ – отец и пр. <p>Прочие родственники (без указания пола):</p> <ul style="list-style-type: none"> • СР – старшие родственники теть, свекровь, теща, дядя и пр. • МР – младшие родственники – зять, невестка, и пр. <p>Друзья</p> <ul style="list-style-type: none"> • БД – близкий друг • БП – близкая подруга • ДР/ПО • АД – ассоциируемый друг (друг друзей, родственников) 	<p>Коллеги</p> <ul style="list-style-type: none"> • КН – коллеги-начальник, командир • КД – коллега-друг или коллега-подруга <p>«Парные» отношения</p> <ul style="list-style-type: none"> • КЛ – клиент (в том числе – покупатель, в транспорте с водителем-кондуктором, на вахте, арендатор квартиры, пациент, родители пациента) • СС – сервис, сервисная служба – все кто оказывают услуги за деньги (продавец, фотограф, квартирная хозяйка, врач) • ССВ – врач • УЧ – ученик • ПР – преподаватель, учитель и пр. 	<p>Группы людей</p> <ul style="list-style-type: none"> • СТ – студент (курсант) • ПП – «попутчик», а также посетитель того же кафе, магазина, поликлиники и пр. <p>Прочее</p> <ul style="list-style-type: none"> • ЗН – просто знакомый, не друг, не клиент, не обслуга, не «партнер» • ЧУ и пр.
--	--	--

Рис. 1. Пример классификации социальных ролей говорящего в корпусе ОРД

Материалом исследования послужили 5 эпизодов звукозаписей одного из информантов корпуса ОРД: **ИЗ8** – мужчина, 53 года, высшее образование, занимается пошивом гидрокостюмов. В табл. 1 приведены метаданные его «речевого дня».

Таблица 1
Материал исследования (ИЗ8)

	Количество эпизодов	Длительность (мин)	Объем подкорпуса (токены)
МУ	3	30	1 533
СС	2	18	1 335
Всего	5	48	2 868

Звукозаписи были расшифрованы с помощью специальной программы для аннотирования аудиофайлов ELAN. Затем текстовые расшифровки были проанализированы в программе eCor, которая позволяет создавать частотные списки (ЧС) словоформ* (см. рис. 2).

* Важно отметить, что в случае токенизации, которая принята при составлении частотных словарей, компоненты единиц типа *потому что, как сказать, так какокажутся* оторваны друг от друга и займут разные строки в списке [Baker, Hardie, McEnery 2006: 159]. Такой формальный подход противоречит задаче исследования: определить, какие единицы являются наиболее употребительными в рамках повседневного общения, поэтому была проведена ручная обработка готовых результатов: все многокомпонентные единицы восстановлены с помощью конкорданса (*как бы, то есть* и др.).

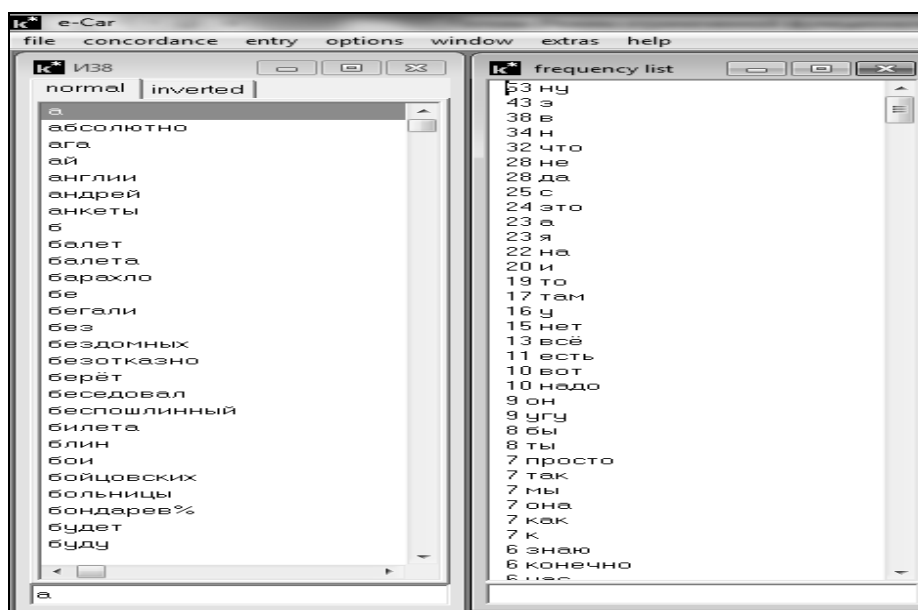


Рис. 2. Фрагмент частотного списка для ИЗ8 в программе eCar

Отличительной чертой частотного списка ИЗ8в роли МУ (см. табл. 2) можно считать появление единицы угу (ранг 26,5), которая и в процентном соотношении (0,6%), отличается от данных ЧС роли СС (0,3%). Единица используется для обозначения некоторых явлений «неканонической фонетики», в данном случае для обозначения утвердительного междометия, произносимого с закрытым ртом (Asinovskyetal. 2009). Это явление уместно в диалоге неформального характера, степень эксплицитности которого снижена в рамках симметричных отношений: «чем интимнее отношения, тем менее эксплицитна речь каждого из участников, тем ярче проявляется тенденция к свертыванию высказываний и замене языковых единиц элементами ситуации, а также жестами, мимикой, телодвижениями и т. п. (Беликов, Крысин 2001: 202).

Таблица 2
Верхняя зона частотного списка ИЗ8в роли мужа

Ранг	Токены	N	%	Ранг	Токены	N	%
1	//	222	14,4	31,5	она	7	0,5
2	?	54	3,5	36,5	есть	6	0,4
3	ну	53	3,4	36,5	не знаю	6	0,4
4	*Н	34	2,2	36,5	конечно	6	0,4
5	()	29	1,9	36,5	нас	6	0,4
6	да	28	1,8	42	то есть	5	0,3
7	в	26	1,7	42	*К	5	0,3
8	это	24	1,6	42	хочу	5	0,3
9	а	23	1,5	42	меня	5	0,3
9	я	23	1,5	42	сейчас	5	0,3
11	не	22	1,4	42	было	5	0,3
11	на	22	1,4	42	значит	5	0,3
13	и	20	1,3	54,5	то (частица)	4	0,3

14	(э-э)	18	1,2	54,5	бы	4	0,3
15	там	17	1,1	54,5	как бы	4	0,3
16	у	16	1,0	54,5	может	4	0,3
17	что	15	1,0	54,5	думаю	4	0,3
17	нет	15	1,0	54,5	Хельсинки	4	0,3
19	(...)	14	0,9	54,5	же	4	0,3
20	*С	13	0,8	54,5	видимо	4	0,3
20	всё	13	0,8	54,5	ей	4	0,3
22	с	12	0,8	54,5	наверно	4	0,3
22	что (QПроп*)	12	0,8	54,5	стало	4	0,3
24	вот	10	0,6	54,5	тоже	4	0,3
24	надо	10	0,6	54,5	они	4	0,3
26	он	9	0,6	54,5	сел	4	0,3
26	угу	9	0,6	54,5	сегодня	4	0,3
28	ты	8	0,5	54,5	три	4	0,3
31,5	*В	7	0,5	54,5	в общем	4	0,3
31,5	(э)	7	0,5	54,5	уже	4	0,3
31,5	просто	7	0,5	73,5	как	3	0,2
31,5	так	7	0,5	73,5	что-то	3	0,2
31,5	мы	7	0,5	73,5	чемоданом	3	0,2

Стремлением к свертыванию высказываний может быть обусловлено и бóльшим количеством фраз (222) в речи И38 в роли МУ, чем в роли СС (159), при разнице между объемами подкорпусов обеих ролей в 197 слов (см. табл. 1).

Обращает на себя внимание частота употребления паралингвистических элементов (ПЭ), которые отражены в расшифровках с помощью специальных знаков: *С – смех, *В – вздох, *К – кашель. ПЭ могут быть отнесены к потенциальным лингвистическим коррелятам социальных ролей, которые помогут обнаружить специфику построения речи говорящего.

В верхней зоне ЧС на материале эпизодов, в которых И38 выступал в роли СС (см. табл. 3) широко представлены хезитационные явления: вербальные хезитативы и паузы хезитаций [3].

Таблица 3
Верхняя зона частотного списка И38 в роли службы сервиса

Ранг	Токены	N	%	Ранг	Токены	N	%
1	//	159	11,8	34,5	всё	6	0,4
2	?	49	3,6	34,5	этот	6	0,4
3	*Н	69	5,1	34,5	тебе	6	0,4
4	ну	50	3,7	37	это самое	5	0,4
5	я	34	2,5	43	потому что	5	0,4
6	да	29	2,2	43	давайте	5	0,4
7	вот	24	1,8	43	понимаешь	5	0,4
8	это	23	1,7	43	давай	5	0,4
9	()	22	1,6	43	же	5	0,4

10	не	21	1,6	43	тебя	5	0,4
11	а	19	1,4	43	меня	5	0,4
12	у	18	1,3	43	или	5	0,4
13	(э-э)	15	1,1	43	к	5	0,4
15,5	(...)	14	1,0	43	тоже	5	0,4
15,5	нет	14	1,0	43	тогда	5	0,4
15,5	на	14	1,0	57,5	есть	4	0,3
15,5	и	14	1,0	57,5	конечно	4	0,3
18,5	что (QPron)	12	0,9	57,5	эти	4	0,3
18,5	с	12	0,9	57,5	вопрос	4	0,3
20,5	(э)	11	0,8	57,5	если	4	0,3
20,5	так	11	0,8	57,5	хочу	4	0,3
22,5	в	9	0,7	57,5	пять	4	0,3
22,5	просто	9	0,7	57,5	сказать	4	0,3
26	как	8	0,6	57,5	тут	4	0,3
26	там	8	0,6	57,5	шапки	4	0,3
26	ты	8	0,6	57,5	угу	4	0,3
26	мы	8	0,6	57,5	вы	4	0,3
26	надо	8	0,6	57,5	них	4	0,3
30,5	как бы	7	0,5	57,5	их	4	0,3
30,5	по	7	0,5	57,5	сейчас	4	0,3
30,5	они	7	0,5	77	что	3	0,2
30,5	но	7	0,5	77	тридцать	3	0,2
34,5	в/на само м деле	6	0,4	77	вас	3	0,2

Стоит отметить, что в ЧС роли СС ранги всех хезитаций выше, чем в роли МУ, существует разница и в процентном соотношении (см. табл. 4).

Таблица 4
Соотношение вербальных хезитативов и пауз хезитаций
в речи ИЗ8 в разных социальных ролях

ПХ/ВХ	МУ (%)	СС (%)
()	1,9	1,6
(...)	0,9	1
(э)	0,5	0,8
(э-э)	1,2	1,1
вот	0,6	1,8
как бы	0,3	0,5
просто	0,5	0,7
Всего	5,9	7,5

Дальнейший акустический анализ позволит выявить возможные качественные отличия хезитационных явлений в разных социальных ролях одного говорящего.

В случае общения в роли СС в верхнюю зону ЧС попадает так называемый контактный глагол *понимаешь*, который служит для привлечения внимания собеседника [6]. В такой же роли выступают единицы *давай/давайте*, которые можно описать как «гортативы» [7, с. 96] или как «типы иницирующих реплик, которые призваны вызывать вербальную реакцию каузируемого слушающего (это может быть подтверждение того, что слушающий не возражает против совершения действия или, напротив, выражение несогласия с совершением действия)» [2, с. 73]:

- *слушайте/давайте* или () в следующую субботу или в воскресенье (И38).

Такое функционирование глаголов в диалоге отсылает к теории вежливости П. Браун и С. Левинсона, описанной в статье «Politeness: Some Universal in Language Usage» [13]. Авторы выделяют категорию социального лица, связанную с коммуникативным взаимодействием в том числе. Человек стремится принадлежать к сообществу и получать одобрение от членов социума (позитивное лицо), так и обладать свободой и независимостью (негативное лицо). С этой точки зрения сама грамматическая форма глагола *понимаешь* направлена на установление контакта и поддержание позитивного лица собеседника, в то время как императивные формы относятся, скорее, к угрозам негативному лицу слушающего [13, с. 65–68].

Единица *это самое* (ранг 37) встречается в верхней зоне ЧС только в роли СС. По функции она близка к хезитативной, но выступает и как маркер поиска, что может обозначать внимание к сказанному со стороны говорящего, возможно, неосознанное.

Обнаруженные особенности верхней зоны ЧС информанта 38 в роли СС позволяют предположить, что такого рода единицы могут стать искомыми лингвистическими коррелятами, которые маркируют социальную роль.

Применение подобной методики анализа к большему объему материала позволит подтвердить это предположение. Однако нужно учитывать, что некоторые явления остаются универсальными для речи мужчин и не будут зависеть от социальной роли в целом: единицы *ну, да, вот***, *а, не, это* располагаются в первой десятке ЧС, по данным аннотированного подкорпуса ОРД (см. рис. 3).

Эти единицы выполняют в устной речи прагматические функции, организуют дискурс. Их позиции в разных ЧС подтверждают их распространенность и значимость в рамках устной коммуникации.

Составление частотных списков для разных СР одного и того же говорящего позволяет обнаружить специфичные черты, которые, несмотря на относительно низкие количественные показатели, могут стать искомыми лингвистическими коррелятами социальных ролей.

** Интересно наблюдение Т. Ю. Шерстиновой, согласно которому единица *вот* чаще употребляется в официальной речи [14], что можно соотнести со случаем асимметричных социальных отношений и определенными социальными ролями. Материал настоящего исследования фиксирует значительную разницу между рангом единицы *вот* в роли МУ (0,6) и в роли СС (1,8).

Ранг	Слово	Абс. кол-во	%
1	ну	2062	2,47
2	я	2000	2,40
3	не	1898	2,28
4	вот	1887	2,26
5	в	1669	2,00
6	да	1610	1,93
7	а	1548	1,86
8	и	1530	1,84
9	там	1456	1,75
10	это	1338	1,60
11	что	1335	1,60
12	на	1113	1,33

Рис. 3. Наиболее употребительные слова мужской устной речи [8, с. 112]

Наиболее представительными в этом аспекте оказались данные ЧС для социальной роли службы сервиса (асимметричные социальные отношения), что подтверждает идею немаркированности ситуаций неформального общения [5, с. 430; 14].

С помощью методов акустического и дискурсивного анализа, а также применения теории речевых актов можно детально охарактеризовать полученные данные и определить те категории, значимость которых в той или иной социальной роли будет определять ее специфику.

Литература

1. Беликов В. И., Крысин Л. П. Социоллингвистика. – М.: РГГУ, 2001. – 315 с.
2. Блинова О. В. Побудительная реплика как элемент устного диалога: случай императива 1Р1 // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной междунар. конф. «Диалог». Вып. 16 (23). Т. 2. – М.: Изд-во РГГУ, 2017. – С. 63–74.
3. Звуковой корпус как материал для анализа русской речи. Ч. 1: Чтение. Пересказ. Описание: коллект. монография / отв. ред. Н. В. Богданова-Бегларян. – СПб.: Филол. факул-т СПбГУ, 2013. – 532 с.
4. Кон И. С. Социология личности. – М.: Политиздат, 1967. – 383 с.
5. Крысин Л. П. Русское слово, свое и чужое: Исследования по современному русскому и социоллингвистике. – М.: Языки славянской культуры, 2004. – 888 с.
6. Маслова Е. Р. *Знаешь и понимаешь* как вербальные хезитативы: сходства и различия // Русская филология. 26: сб. науч. работ молодых филологов. – Тарту: Тартуский ун-т, 2015. – С. 275–279.
7. Подлесская В. И. Русские глаголы *дать/давать*: от прямых употреблений к грамматикализованным // Вопросы языкознания. – 2005. – № 2. – С. 89–103.
8. Русский язык повседневного общения: особенности функционирования в разных социальных группах: коллект. монография / отв. ред. Н. В. Богданова-Бегларян. – СПб.: ЛАЙКА, 2016. – 244с.
9. Чеботникова Т. А. Социальная роль и речевое поведение личности: инвариант и варианты [Электронный ресурс] – URL: <https://cyberleninka.ru/article/n/sotsialnaya-rol-i-rechevoe-povedenie-lichnosti-invariant-i-varianty> (дата обращения: 11.12.2017).
10. Швейцер А. Д. Современная социоллингвистика: Теория, проблемы, методы / 2-е изд. – М.: Книжный дом «ЛИБРОКОМ», 2009. – 176 с.

11. Asinovsky A., Asinovsky A., Bogdanova N., Rusakova M., Ryko A., Stepanova S., Sherstinova T. The ORD Speech Corpus of Russian Everyday Communication «One Speaker's Day»: Creation Principles and Annotation / Matoušek V., Mautner P. (eds.). TSD 2009. LNAI, vol. 57292009. – Berlin-Heidelberg: Springer, 2009. – Pp. 250–257.

12. Baker P., Hardie A., McEnery T. A Glossary of Corpus Linguistics. – Edinburgh: Edinburgh Univ. Press. Journal articles, 2006. – 192 p.

13. Brown P., Levinson S. Politeness: Some Universals in Language Usage. – Cambridge: Cambridge Univ. Press, 1987. – 358 p.

14. Sherstinova T. Macro Episodes of Russian Everyday Oral Communication: towards Pragmatic Annotation of the ORD Speech Corpus. *SPECOM 2015*, Lecture Notes in Artificial Intelligence, LNAI, vol. 9319. – Springer, Switzerland, 2015. – Pp. 268–276.

FREQUENCY LISTS AS A MEAN OF SEARCHING LINGUISTIC CORRELATES OF SOCIAL ROLES OF THE SPEAKER

Tatiana I. Popova

Research Assistant,

St. Petersburg State University

7/9 University Emb., St. Petersburg 199034, Russia

E-mail: tipopova13@gmail.com

The paper deals with analysis of frequency lists, compiled on the basis of speech transcripts of one informant's audio recordings from the ORD Speech corpus. Study material consists of 5 episodes in which the informant acts in social roles, opposing each other on the principle of asymmetric social relations: the men's role as a *husband* when communicating with his wife and when communicating with clients, identified in the ORD Speech corpus as a *service*. The subject of the study are the most commonly used units from the upper band of the frequency list. They can be interpreted as linguistic correlates of the speaker's social role due to their specificity, but the special subcode, which they create, is described in linguistics only in the theoretical aspect. In this case, the situation of formal communication turns out to be marked. Further analysis of the extracted units will allow to determine what is the significance of the social role in communication and what is the position of the communicative situation when code-switching behavior.

Keywords: colloquial Russian; corpus; sociolinguistics; social role; communicative situation; linguistic correlates; frequency lists; discourse analysis; politeness theory; communication strategies