***Н.В. Богданова-Бегларян***

**МАРКЕРЫ-КСЕНОПОКАЗАТЕЛИ В РУССКОЙ ПОВСЕДНЕВНОЙ РЕЧИ: АННОТИРОВАНИЕ РЕЧЕВОГО КОРПУСА, ТИПОЛОГИЯ И КОЛИЧЕСТВЕННЫЕ ДАННЫЕ[[1]](#footnote-1)**

**Аннотация.** В статье предлагаются результаты анализа аннотированных выборок из двух речевых корпусов. Разметка проводилась для выявления в материале прагматических маркеров; описанию, систематизации и количественной характеристике подверглись маркеры одного из десяти выделенных классов – ксенопоказатели. В статье приведена их типология и данные об их реализации в речи разных говорящих, в разных коммуникативных ситуациях.

**Ключевые слова**: повседневная речь, звуковой корпус, аннотирование корпуса, прагматикализация, прагматический маркер, ксенопоказатель.

1. **Введение**

*Ксенопоказатели* рассматриваются в статье как класс прагматических маркеров (ПМ). Последние возникают в устной речи в результате последовательно действующих процессов *грамматикализации* и *прагматикализации*, в ходе которых единицы сначала меняют свой грамматический статус, а затем и вовсе практически утрачивают и лексическое, и грамматическое значение и приобретают прагматическое, т.е. определенную функцию в дискурсе. Для маркеров-ксенопоказателей (ПМК) такой функцией становится указание на «присутствие Другого» [Арутюнова 2000: 448]: это и чужая речь, и своя собственная, сказанная ранее или только планируемая на будущее, собственные или чужие мысли и даже речевое, «говорящее», поведение другого человека, его реакции и т.д. «Классические» ксенопоказатели, возникшие из глаголов, – *мол, де, дескать* – остановились в своем развитии на этапе грамматикализации и сегодня описываются как частицы. «Новые» ксенопоказатели чаще проходят другой путь грамматикализации: через вводные единицы – к ПМК. Их перечень в современной речи очень обширен и имеет столь же обширный набор «протоединиц»: вводные единицы, глаголы, наречия, местоимения, союзы. В результате многие ПМК сохранили ограниченную грамматическую изменяемость (*грю/гришь/грит/грим; такой/ая/ие*), что и позволяет числить их в классе не частиц, но – прагматических маркеров (ср.: [Fraser 1996]).

1. **Источники материал и методика исследования**

Основными источниками материала для настоящего исследования стали два речевых корпуса, созданных в СПбГУ: корпус повседневной русской речи «Один речевой день» (ОРД) и корпус русской монологической речи «Сбалансированная аннотированная текстотека» (САТ) [см. о них: [Звуковой корпус… 2013; Русский язык… 2016]. В целях получения максимально полного инвентаря тех функциональных единиц, которые использует человек в ходе речепорождения, а также в целях описания их функционирования в речи разных говорящих и в разных коммуникативных ситуациях корпусный материал был проаннотирован с выявлением в нем прагматических маркеров. Такая разметка оказалась весьма затруднена тем обстоятельством, что ПМ внешне ничем не отличаются от своих «протоединиц» и лишь в контексте реализуют свой новый статус. В результате ПМ, как и многие дискурсивные слова, «можно изучать только через их употребление» [Дискурсивные слова… 1998: 8‑10]. Кроме того, ПМ часто полифункциональны и склонны выстраиваться в цепочки однофункциональных единиц (о проблемах корпусной разметки ПМ и способах их решения см.: [Bogdanova-Beglarian et al. 2018, 2020; Богданова-Бегларян и др. 2019 а, б]). Дополнительным источником материала для наблюдений и выводов стал устный подкорпус (УП) НКРЯ.

На этом материале был получен, в числе прочего, относительно полный набор ПМК и статистика их встречаемости в речи разных говорящих и в разных коммуникативных ситуациях.

1. **Ксенопоказатели устной речи в зеркале статистики**

Исследование произведено на аннотированных выборках ОРД (195 эпизодов «речевых дней» 104 информантов, более 300 тыс. словоупотреблений) и САТ (монологи разного типа от 34 информантов, более 50 тыс. словоупотреблений). Общий словник ПМ, полученных на этом материале, составил 60 единиц. Их типология включает 10 разновидностей: хезитативы (Х), рефлексивы (Ф), метакоммуникативы (М), разграничители (Г), ксенопоказатели (К), аппроксиматоры (А), дейктические (Д) и ритмообразующие (Р) маркеры, маркеры самокоррекции (С) и заместители (З) [Прагматические маркеры… 2021].

В общем частотном списке функций ПМ ксенопоказатели имеют ранг 5 и уступают по встречаемости в нашей речи только хезитативам, разграничителям, метакоммуникативам, а также смешанной группе ГХ. Такая ситуация свойственна и большинству социолектов, выделенных по признакам пола, возраста, образования и статуса говорящих, уровня их речевой компетенции (УРК) и рода деятельности. Обнаружились лишь некоторые исключения. Больше всего ксенопоказателей (ранг 2) оказалось в речи работников службы сервиса, а также учащихся и руководителей, гуманитариев и людей с незаконченным высшим образованием (ранг 3). Совсем их не встретилось в речи детей, что наводит на мысль, что способность употреблять такие маркеры приобретается только с речевым опытом.

Тот же ранг 5 имеют ксенопоказатели и в речи экстравертов, тогда как в топ-5 частотного списка функций в речи интровертов ПМК вообще не попали. Наиболее частотен в речи экстравертов оказался маркер ГОВОРИТ (в его редуцированной форме *грит*) – третье место после ВОТ и ТАМ, с частотой 6,05 %.

Учет типа и места коммуникации показал, что реже всего ПМК используются в публичной и учебной речи (ранги соответственно 8 и 11), в казарме и в кафе/ресторанах (ранги 11 и 10). Чаще всего – в разговорах в поликлинике (ранг 2).

Из конкретных ксенопоказателей наиболее частотен маркер *говорит* – в разных его редуцированных формах: *грит* (1116 ipm), *грю* (669), *грят* (47), *гришь* (10). Частотность остальных ПМ, как инвариантов, так и структурных разновидностей, установить затруднительно в силу их высокой полифункциональности. Конкретную функцию маркера можно установить лишь в контексте.

1. **Опыт систематизации маркеров-ксенопоказателей**

Систематизировать ксенопоказатели в корпусном материале можно, например, с опорой на их прототип (условная чужая речь в контекстах подчеркнута, ПМК выделен).

1. Вводная единица → ПМК:

* *Вы пришли в булочную и Вас там у кассы облаяли / потому что* ***видите ли*** *у Вас крупная бумага / а у них сдачи нет* (УП);
* *ну или надо посмотреть где арен… в аренду парики сдают / ну там же не будет* ***так сказать*** *ой нам надо всего на десять минут дайте нам его / за сто рублей и они скажут идите в ж-ж\*пу может быть* (ОРД);
* *он такой / понял* ***короче*** */ по... посмеялся / поулыбался // \*Н без темы // \*П ну вот / я помогу* ***короче*** (ОРД)[[2]](#footnote-2).

1. Союз → ПМК:

* *какой будет / прогноз* *? \*П* # *да* ***вроде*** *у нас теплеет* (ОРД);
* *Ну / она* [нрзб] *когда отвечаю / то в окно смотрит там / то лицо такое делает /* ***типа*** *я тупая* (УП);
* *Если взять так / тот же Билл Гейтс /* ***якобы*** *образовавший компанию «Майкрософт» / который самый богатый человек…* (УП).

1. Глагол → ПМК:

* *ну вот // \*В я говорю* ***говорю*** */ я* ***говорю*** */ барышня !* ***говорю*** */ какая вы молодца* ***говорю*** *// правильно / () а то / я некрасивая ! сделайте меня покрасивше* ***говорю***(ОРД);
* *и(:) Наталья\_Георгиевна% мне (э) всё говорила /* ***знаешь*** */ их тогда никто не обойдёт на(:) этапе поставки* (ОРД);
* *и он блин мне написал* ***слушай*** *сорри не могу* (ОРД).

1. Местоимения → ПМК:

* *я вчера их встречаю / на улице / ну в «Пятёрочку»$ / я шла как раз в магазин / а был вечер / полдесятого // я****такая*** */ о-о / вы приехали // а у вас завтра занятия будут? / они* ***такие*** */ будут / я говорю / да-а / не повезло мне // а что такое ? // а у меня завтра у вас три ... два семинара* (ОРД);
* *Её / на сутки позже / принесли на кормление… Его оживляли тоже. Бук… буквально оживляли. В общем я там взволновалась / но они / очень боялись что у меня там молоко пропадёт /* ***то-сё*** */ «не волнуйтесь / не волнуйтесь»* (УП);
* *Я ж и говорю, вам бы все читать, а если вам живого человека дают, так вы,* ***это самое****, зарежет меня живой человек* (ОП) (единственный пример из основного подкорпуса НКРЯ).

1. Наречие → ПМК:

* *а-а \*П э судья на Колю%* ***так*** */ тш(:) ! \*П вы не на базаре* (ОРД).

1. Комбинации разных частей речи → ПМК:

* *я подошёл к тренеру и говорю ему* ***так и так******вот*** *я занимался водным поло / энное количество лет / я хочу попасть в команду* (САТ);
* *вот один из них спросил таким / \*К \*П так / это просто со странным немножко (...) отношением и взглядом / \*П а зачем ? \*П ну* ***вроде того что*** *их интересует только живая / (...) музыка // живой фольклор* (ОРД);
* *и тут звонок в дверь // стоит этот мужик // \*П* ***типа того что*** *блин / давайте общаться !* (ОРД);
* *и она как на нас налетела !* ***вот******там*** *ты-ты-ты-ты-ты-ты / да мы алкаши там / ну что-то там такое / я не помню* (ОРД);
* *Вот / пишет мне* ***типа*** *ну вот* ***так и так / типа*** *перевод / я говорю / "Ну деньги сначала / да / там все дела"* (УП).

Возможны, вероятно, и иные способы систематизации ксенопоказателей: проаннотированный корпусный материал предоставляет такие возможности.

1. **Некоторые выводы**

Анализ материала показал, что круг маркеров-ксенопоказателей в повседневной речи очень широк и значительно превышает известные из литературы перечни (см., например: [Левонтина 2020]. Как правило, все ПМК полифункциональны (исключение – *якобы*:только К – см.: [Прагматические маркеры… 2021]) и склонны выстраиваться в цепочки. Анализ демонстрирует «отрыв» ПМК от глаголов говорения (ср.: *де, дескать, мол*) и явное преобладание ПМК в диалоге по сравнению с монологом: ОРД –2336 ipm, ранг 5; САТ – 420 ipm, ранг 8.

Полный перечень ПМК может быть полезен в самых разных прикладных аспектах лингвистики: преподавание РКИ, практика перевода, лингвистическая экспертиза, автоматическая обработка речи, лингвокриминалистика.

**Литература**

1. Арутюнова Н.Д. (2000), Показатели чужой речи *де, дескать, мол*. *Язык о языке: сб. статей*. М.: Языки русской культуры, с. 437‑452.
2. Богданова-Бегларян Н.В., Блинова О.В. , Шерстинова Т.Ю., Зайдес К.Д., Попова Т.И*.* (2019а), **Корпус естественной речи: проблемы ручного аннотирования прагматических маркеров и пути их решения.** *Анализ разговорной русской речи (АР3-2019): Труды восьмого междисциплинарного семинара*). СПб.: Политехника-принт, СПб., с. 5‑10**.**
3. Богданова-Бегларян Н.В., Блинова О.В. , Мартыненко Г.Я., Шерстинова Т.Ю., Зайдес К.Д., Попова Т.И. (2019б), Аннотирование прагматических маркеров в русском речевом корпусе: проблемы, поиски, решения, результаты. *Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной международной конференции «Диалог».* Вып. 18 (25), с. 72-85.
4. Дискурсивные слова русского языка: Опыт контекстно-семантического описания (1998). М.: Метатекст, 448 с.
5. Звуковой корпус как материал для анализа русской речи. Коллективная монография. Ч. 1. Чтение. Пересказ. Описание (2013). СПб.: Филологический ф-т СПбГУ, 532 с.
6. Левонтина И.Б.(2020), Об арсенале ксенопоказателей в русском языке. *Вопросы языкознания*. № 3, с. 52–77.
7. Прагматические маркеры русской повседневной речи: Словарь-монография (2021). СПб.: Нестор-История. В печати*.*
8. Русский язык повседневного общения: особенности функционирования в разных социальных группах. Коллективная монография. (2016). СПб.: ЛАЙКА, 244 с.
9. Bogdanova-Beglarian, N., Blinova, O., Martynenko, G., Sherstinova. T., Zaides, K.(2018), Pragmatic Markers in Russian Spoken Speech: an Experience of Systematization and Annotation for the Improvement of NLP Tasks. *Proceedings of the FRUCT’23*. Bologna, Italy. FRUCT Oy, Finland, pp. 69-77.
10. Bogdanova-Beglarian, N., Blinova O., Sherstinova T., Gorbunova D, Zaides K., Popova T.(2020), Pragmatic Markers in Dialogue and Monologue: Difficulties of Identification and Typical Formation Models. *SPECOM 2020. Lecture Notes in Artificial Intelligence, LNAI*, vol. 12335. Springer, Switzerland, pp. 68-78.
11. Fraser, B. (1996), Pragmatic Markers. *Pragmatics*. Vol. 6 (2), pp. 167–190.

**References**

1. Arutyunova, N.D. (2000), Pokazateli chuzhoj rechi *de, deskat', mol*.Jаzyk o jazyke: sb. statej. Moscow, pp. 437-452.
2. Bogdanova-Beglarian, N., Blinova O., Sherstinova T., Zaides K., Popova T.(2019a), Korpus jestestvennoj rechi: problemy ruchno-go annotirovania pragmaticheskikh markerov i puti ikh reshenia. Analiz razgovornoj russkoj rechi (AR3-2019): Trudy vos'mogo mezhdis. seminara. St Petersburg, pp. 5-10.
3. Bogdanova-Beglarian, N., Blinova O., Martynenko, G., Sherstinova T., Zaides K., Popova T.(2019b), Annotirovanie pragmati-cheskikh markerov v russkom rechevom korpuse: problemy, po-iski, reshenia, rezul'taty. Kompjuternaya lingvistika i intellektual'nye tekhnologii: Po m-lam jezhegod. mezhdun. konf. «Dialog». Iss. 18 (25). Moscow, pp. 72-85.
4. Diskursivnye slova russkogo jazyka: Opyt kontekstno-semanticheskogo opisania (1998). Moscow, 280 p.
5. Zvukovoj korpus kak material dl’a analiza russkoj rechi. Kol. monogr. CH. 1. Chtenie. Pereskaz. Opisaniye (2013). St Petersburg, 532 p.
6. Levontina I.B. (2020), Ob arsenale ksenopokazatelej v russkom jazyke. Voprosy jazykoznania. No. 3, pp. 52-77.
7. Pragmaticheskie markeryrusskoj povsednevnoj rechi: slovar’-monografia (2021). St Petersburg. In Print.
8. Russkij jazykpovsednevnogo obshchenia: osobennosti funktsi-onirovania v raznykh sotsial'nykh gruppakh. Col. monogr. (2016). St Petersburg, 244 p.
9. Bogdanova-Beglarian, N., Blinova, O., Martynenko, G., Sherstinova. T., Zaides, K.(2018), Pragmatic Markers in Russian Spoken Speech: an Experience of Systematization and Annotation for the Improvement of NLP Tasks. Proceedings of the FRUCT’23. Bologna, Italy. FRUCT Oy, Finland, pp. 69-77.
10. Bogdanova-Beglarian, N., Blinova O., Sherstinova T., Gorbuno-va D, Zaides K., Popova T.(2020), Pragmatic Markers in Dia-logue and Monologue: Difficulties of Identification and Typical Formation Models. SPECOM 2020. Lecture Notes in Artificial Intelligence, LNAI, vol. 12335. Springer, Switzerland, pp. 68-78.
11. Fraser, B. (1996), Pragmatic Markers. Pragmatics. Vol. 6 (2), pp. 167–190.

**‘XENO’-MARKERS IN RUSSIAN EVERYDAY SPEECH: ANNOTATION OF THE SPEECH CORPUS, TYPOLOGY AND QUANTITATIVE DATA**

N. . Bogdanova-Beglarian

**Abstract.** The article proposes the results of the analysis of annotated samples from two speech corpora created at St. Petersburg State University: the corpus of everyday Russian speech "One Speaker’s Day" (ORD) and the corpus of Russian monologic speech "Balanced Annotated Text Library" (SAT). The corpus annotation was carried out to identify pragmatic markers (PM) in the material. The markers of one of the ten identified PM classes were described – ‘xeno’-markers (PMK): *grit/gr’u, takoi/aya/ie, tipa (togo chto), vrode (togo chto), znaesh’, slushaj, tak*, etc. The grammatical specificity of these units allows us to consider them not as particles but as pragmatic markers. The article provides a typology of PMK and quantitative data of their use in different speakers’ discourse and diverse communicative situations.

**Keywords:** everyday speech, speech corpus, corpus annotation, pragmatikalization. pragmatic marker, ‘xeno’-marker.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Богданова-Бегларян Наталья Викторовна**

Санкт-Петербургский государственный университет (Россия)

**Bogdanova-Beglarian Natalia**

Saint-Petersburg State University (Russia)

n.bogdanova@spbu.ru

1. **Исследование проведено при финансовой поддержке гранта Санкт-Петербургского государственного университета (проект № 75254082 «Моделирование коммуникативного поведения жителей российского мегаполиса в социально-ре-чевом и прагматическом аспектах с привлечением методов искусственного ин-теллекта»).** [↑](#footnote-ref-1)
2. Об особенностях дискурсивной транскрипции материала ОРД см.: [Русский язык… 2016: 242-243]. [↑](#footnote-ref-2)