

RU

Дискурсивная транскрипция устной речи диалогического типа

Михалёва О. Л., Чекмез У. Э.

Аннотация. Цель исследования - обосновать необходимость использования особой системы аннотирования дискурса, учитывающей специфические для устной речи диалогического типа явления: наложения, паузы внутри речевой партии говорящего, маркеры обратной связи; рассмотреть варианты фиксации этих явлений. Научная новизна обуславливается тем, что предложенная система, во-первых, дополняет рядом условных обозначений уже имеющиеся применительно к дискурсу диалогического типа; во-вторых, разработка этой системы позволяет осмыслить специфические характеристики диалогического дискурса как особой разновидности устного дискурса. В результате предложен вариант дискурсивной транскрипции, позволяющий с высокой степенью точности фиксировать явления, релевантные для диалогической речи.

EN

Discourse Transcription of Oral Dialogic Speech

Mikhaleva O. L., Chekmez U. E.

Abstract. The research objectives are as follows: to justify reasonability of using a discourse transcription system that takes into account peculiarities of dialogic speech (overlays, speech pauses, feedback markers), to analyse possible forms of their written representation. Scientific originality of the study lies in the following. Firstly, the proposed system expands the existing transcription systems by adding special symbols peculiar for dialogic discourse; secondly, the developed system makes it possible to better understand dialogic discourse as a special type of oral discourse. As a result, the authors propose the discourse transcription system which allows more accurate representation of dialogic speech peculiarities.

Введение

Исследователи, изучающие устную речь, неизбежно сталкиваются с проблемой её фиксации. Существуют как сложности технического характера, связанные со сбором материала (наличие или отсутствие специальных технических средств для записи устной речи и их качество, количество участников общения (чем большее количество человек участвует в разговоре, тем сложнее зафиксировать речевую партию каждого участника), обстановка, в которой происходит запись и т.п.), так и содержательные сложности, связанные с переводом устной речи в графический текст (любой, кто в качестве объекта исследования выбирает устную речь, неизбежно сталкивается с проблемой, касающейся выбора того или иного способа её аннотирования). Это обусловлено тем, что «устный дискурс, по своей природе текучий и континуальный, не может быть предметом научного рассмотрения, если он представлен лишь в форме звука. Чтобы всерьёз заниматься анализом дискурсивного процесса, он должен быть репрезентирован в графической форме, представлен как объект, который можно рассматривать с разных точек зрения» [7, с. 17].

Актуальность темы исследования продиктована как отсутствием общепринятой системы аннотирования устного дискурса, так и необходимостью корректировки существующей системы аннотирования диалогического дискурса, используемой сейчас при конвертации устного диалогического дискурса.

Для достижения указанной цели исследования необходимо решить следующие задачи: во-первых, обобщив результаты современных научных исследований, посвященных описанию устной речи, критически осмыслить выдвинутые положения в преломлении к анализу устного дискурса диалогического типа; во-вторых, разработать и ввести ряд условных обозначений, способных максимально точно зафиксировать явления, релевантные для аннотирования устной диалогической речи.

В ходе исследования в качестве основного метода использовался индуктивный метод, включающий непосредственное наблюдение, анализ, сопоставление, классификацию речевых и коммуникативных фактов и позволяющий тем самым обнаружить закономерности организации устной диалогической речи и разработать основные положения аннотирования устного дискурса этого типа.

Теоретической базой исследования послужили труды Е. А. Земской, Л. А. Капанадзе, М. В. Китайгородской, Н. Н. Розановой [11; 12], посвященные описанию русской разговорной речи; работы А. А. Кибрика, Н. А. Коротаева, В. И. Подлесской, О. В. Федоровой, А. А. Евдокимовой, С. И. Мазуриной [5; 7-10; 15], в которых предложена система аннотирования коммуникативных каналов, ориентированная на комплексное описание устного дискурса. Кроме того, учитывались работы Н. В. Богдановой-Бегларян, О. В. Блиновой, Г. Я. Мартыненко, Т. Ю. Шерстиновой, К. Д. Зайдес, Т. И. Поповой [1; 3; 4], в которых описываются основные типы спонтанной монологической речи; монография И. Н. Борисовой [2], посвященная русскому разговорному диалогу, и работы Т. Е. Янко [13; 14], представляющие собой исследования стратегий, используемых в русской устной речи.

Практическая значимость работы состоит в возможности использования принципов транскрипции для аннотирования диалогического дискурса при создании устного корпуса. Материалы исследования могут быть использованы в практике преподавания ряда теоретических и прикладных лингвистических дисциплин, таких как дискурс-анализ, теория коммуникации, компьютерные методы в лингвистике и др. Полученные данные могут найти применение в процессе учебно-методической деятельности при создании учебников, учебных пособий, направленных на изучение устного диалогического дискурса.

Введение в проблему

Устный дискурс может быть зафиксирован на письме различными способами, например, с использованием системы записи, существующей для письменной формы речи, как это делается при оформлении стенограммы устного доклада, заседания, интервью и т.п. Ещё в первых трудах, посвященных изучению устной речи, отмечалось, что «многие исследователи литературной разговорной речи, публикуя фрагменты связного текста, пользуются традиционной пунктуацией» [11, с. 21], однако если устный дискурс проходит обработку по типу стенограммы, то в этом письменном варианте почти целиком стираются специфические черты устной речи. Такой способ графической фиксации материала, используемый, в частности, в 60-х годах XX в., когда на устную форму речи ещё только стали обращать внимание, кажется для того времени вполне объяснимым и даже приемлемым, однако и спустя 50 лет в современных работах встречаем утверждения, сходные по содержанию с высказыванием Е. А. Земской: «Значительный прогресс, достигнутый в последние годы российской корпусной лингвистикой, к сожалению, в весьма малой степени связан с корпусным представлением устных данных. <...> Не менее существенна форма, в которой представлен языковой материал. Обычно это максимально упрощенная графическая нотация, не учитывающая такие центральные для устного модуса феномены, как сегментация речевого потока, интонация, паузация, принципы дискурсивной организации» [8, с. 205]. «Практически во всех работах, посвященных анализу устной речи, отмечается стремление авторов привести устный текст в соответствие с его письменной формой. Письменная фиксация устных текстов неизбежно влечет за собой оформление их по правилам письменной речи, включение аннотации, лишь частично реагирующей на интонационные характеристики» [3, с. 15]. В современных условиях, когда лингвисты имеют возможность использовать различные программы, предназначенные для анализа устной речи, такие как ELAN, Praat, Speech Analyzer и т.п., когда имеется ряд работ (см., например, Кибрик, Подлесская, 2009 [10]; Коротаев, 2019 [7]), в которых сформулированы базовые принципы аннотации, позволяющие отражать релевантные характеристики устного дискурса, такое «приблизительное», неточное представление столь ценного материала не кажется оправданным.

Наличие значительной доли работ, в которых устный дискурс представлен в максимально упрощенной форме, может быть обусловлено, в частности, тем, что «представление устной речи в графическом виде является сложной и нетривиальной научной задачей. Преобразование исходного разворачивающегося во времени дискурса в графический текст, т.е. статический объект, – это аналитическая процедура, подразумевающая большое количество решений» [10, с. 29]. Это может быть обусловлено и тем, что подробное аннотирование устной речи «означает разработку десятков и сотен различных конвенций, которые позволяли бы принимать транскрипционные решения осмысленно и систематично» [Там же, с. 17]. Всё это требует от исследователя высокого уровня компетентности в анализе фонетических явлений различного типа, в том числе – с помощью современного программного обеспечения.

Перед исследователем, избравшим в качестве объекта анализа устную речь, встают разнообразные задачи, от способов решения которых может зависеть степень подробности аннотации устного материала, обеспечивающая получение достоверных результатов исследования. Так, если перед лингвистом стоит задача описать специфическую для монологической устной речи лексику или функционирование в устной речи отдельных грамматических форм, то создание подробной аннотации с фиксацией заполненных/незаполненных пауз и движения тона в таком случае не представляется целесообразным. В другой ситуации оказывается, например, лингвист, перед которым стоит задача описать когнитивные процессы, происходящие в сознании говорящего при пересказе некоторого текста, в этом случае исследователь может не отражать в аннотации маркеры обратной связи, порождаемые слушателем, и может описывать речь говорящего как монологическую, однако это уже может повлиять на валидность результатов исследования, так как степень влияния реакций слушателя на порождение речи основного говорящего ещё недостаточно изучена. Если же перед ученым стоит задача описать функционирование диалогического дискурса, то он должен учесть в своей аннотации максимально возможное количество релевантных для диалогического дискурса характеристик.

Таким образом, если оставить за рамками технические сложности, возникающие при сборе материала, и вопрос об уровне компетентности конкретного исследователя, можно заключить, что выбор того или иного типа аннотации определяется типом устного дискурса (монолог / диалог / полилог, рассказ о жизни / дискуссия / публичное выступление) и – в ещё большей степени – задачами, стоящими перед ученым той или иной области гуманитарного знания.

Обзор существующих принципов аннотации устной речи

Первый опыт аннотирования русской устной речи представлен в работах, выполненных группой учёных под руководством Е. А. Земской. В книге «Русская разговорная речь. Тексты» сформулированы основные теоретические принципы, положенные в основу аннотирования устной речи: «Составители не ставили своей целью полностью передать фонетику разговорной речи, т.е. отказались от представления всех текстов в фонетической транскрипции. Наш материал должен был и в письменной передаче восприниматься без смысловых потерь, форма передачи речи не должна была стать кривым зеркалом, которое представило бы в искаженном виде факты морфологии, словообразования, лексики и особенно синтаксиса РР (разговорной речи. – О. М., У. Ч.). Не менее важную задачу мы видели и в том, чтобы как можно более точно отразить строение связанного текста в РР и строение диалога и полилога» [11, с. 17]. Эти принципы в той или иной мере используются и в современных нотациях. В частности, в названной выше работе аргументировано использование традиционной графики и орфографии для передачи устной речи, описаны принципы выделения словесного ударения, предложено графическое отражение некоторых специфических явлений устной речи, таких как стяженные формы и контрасты темпа внутри одного текста, указано на необходимость фиксировать приблизительное произношение некоторых форм, передающих индивидуальные характеристики говорящего, отмечена важность передачи заполненных и незаполненных пауз хезитации, «растяжек» гласных и согласных (в текстах, представленных в книге, они фиксируются с ориентацией на устойчивые символы художественной литературы). Помимо этого авторы обращают внимание на случаи нейтрализации противопоставления форм в разговорной речи (*Мам/ вот эт[ъ] шла/*), пытаются передать на письме синтаксическое и интонационное членение разговорной речи, «накладки» одних реплик на другие, перебивы, вставки реплик одного говорящего в текст другого и «некоторые явления неканонической фонетики», такие как маркеры обратной связи, которые предлагается передавать «орфограммами» *угу, гм, не-у, м-м, э-э, м?, не-а* [Там же].

Такое количество специфических черт устной речи было выявлено ещё на первых этапах её изучения. Это обусловлено разнообразием материалов, зафиксированных исследователями. Как кажется, данная выборка хорошо отражает характеристики «генеральной совокупности» устных текстов, что и позволило группе Е. А. Земской отметить столько релевантных для устного дискурса особенностей, выявление которых на столь раннем этапе изучения подтверждает определяющую роль этих характеристик для устного дискурса, а также выступает в качестве аргумента в пользу обязательности фиксации этих явлений в аннотации. Несмотря на то что исследователям удалось обнаружить такое количество особенностей устной речи, технические возможности не позволяли фиксировать с достаточной степенью точности многие из них, в частности, авторы отмечают, что тексты хрестоматии не рассчитаны на передачу интонации, хотя уже на тот момент в их корпусе существовало несколько отрывков, размеченных по системе Е. А. Брызгуновой.

Следует отметить, что лишь немногие феномены устной речи, выявленные в начале 1960-х, были исчерпывающе описаны. Так, несмотря на регулярно встречающиеся упоминания о разнообразии маркеров обратной связи (см. например, «а разговорные междометия – все эти *м?, угу, ага* и мн. др., которые столь важны в устной речи, но для передачи которых на письме нет адекватных буквенных средств» [Там же, с. 6-7]; «с точки зрения происхождения релятивы очень разнообразны» [12, с. 92-93]), эти маркеры не были хоть сколько-нибудь полно описаны с точки зрения семантики и функционирования; более того – даже их инвентарь в полном объеме не установлен.

Во многих работах постулируется существование значимых отличий синтаксиса устной речи от синтаксиса письменной, см. например: «Как это уже признано многими исследователями, в разговорной речи реализуется особая система синтаксиса. Обнаружены специфически разговорные: синтаксические модели, особые средства синтаксической связи и синтаксической организации единиц этого уровня в целом. Многие исследователи приходят к выводу о невозможности описывать синтаксис разговорной речи в тех же единицах, что и синтаксис кодифицированного литературного языка. В частности, оказывается практически невозможным разделить разговорный текст на предложения, как они определяются в кодифицированном литературном языке» [11, с. 20]. Между тем не были выявлены и описаны единицы, которые могли бы служить адекватным коррелятом предложению. Так, в аннотациях, предложенных коллективом Е. А. Земской, в качестве единицы членения текстов избирается высказывание, которое оформляется с помощью специальных знаков членения: «//, ! и ? – это знаки завершения высказывания. / – это знак интонационного членения высказывания при его незаконченности» [Там же, с. 21]. Несмотря на выдвигание на первый план диалогичности устной речи, не были описаны причины возникновения таких явлений, как сцепление реплик, изменение стимулов и реакций под влиянием собеседника, оговорки, переспросы, не были сформулированы принципиальные отличия диалогической устной речи от монологической.

Дальнейшее развитие исследований, касающихся изучения устной речи с точки зрения аннотирования дискурса, представляется довольно любопытным. Так, в 2000-х годах вышло несколько работ, посвященных

анализу устной речи: Т. Е. Янко «Коммуникативные стратегии устной речи» [13], И. Н. Борисовой «Русский разговорный диалог. Структура и динамика» [2], «Рассказы о сновидениях: корпусное исследование устного русского дискурса» под ред. А. А. Кибрика и В. И. Подлесской [10]. Обратим внимание на две из них, так как они основывались на устном материале, представляющем собой корпус расшифрованных исследователями записей устной речи: это монография И. Н. Борисовой «Русский разговорный диалог. Структура и динамика» и коллективная монография под ред. А. А. Кибрика и В. И. Подлесской «Рассказы о сновидениях: корпусное исследование устного русского дискурса».

В работе И. Н. Борисовой «расшифровка звучащей речи сделана с опорой на методику оформления разговорных текстов, принятую на кафедре риторики и стилистики русского языка Уральского государственного университета, в целом приемственную по отношению к методике, разработанной коллективом учёных Института русского языка РАН им. В. В. Виноградова» [2, с. 5], то есть аннотация базируется на принципах, изложенных в 1960-х годах коллективом под руководством Е. А. Земской, в то время как в работе «Рассказы о сновидениях» под ред. А. А. Кибрика и В. И. Подлесской представлены результаты разработки новой системы транскрипции устного дискурса, описанию которой посвящено около 60% текста монографии. Позже, в 2010-х, опубликованы монографии петербургского коллектива, работающего под руководством Н. В. Богдановой-Бегларян, под общим названием «Звуковой корпус как материал для анализа русской речи» (2013-2015) [3; 4]. В них не используется способ аннотирования вокального канала дискурса, предложенный в работе «Рассказы о сновидениях», данная группа исследователей придерживается скорее принципов, применявшихся при расшифровке ещё в 1960-х годах коллективом Е. А. Земской.

Специфика аннотирования устного диалогического дискурса

В рамках нашего исследования материал аннотировался в соответствии с принципами, предложенными в монографии «Рассказы о сновидениях». Система, описанная в этой работе, обеспечивает возможность такой аннотации устной речи, при которой сразу осуществляется анализ и классификация некоторых явлений, например, самокоррекции. Принципиальным отличием предложенного способа аннотирования является системность, стремление к фиксации всех релевантных характеристик устной речи, отмеченных выше, а также разработанный комплекс условных обозначений, позволяющих создавать универсальные аннотации. Точная временная привязка и разделение потока речи на элементарные дискурсивные единицы (далее ЭДЕ) обеспечивают возможность оценки предельной длительности некоторых единиц. Несмотря на все достоинства предложенной системы, она не оказалась достаточно полной для аннотирования устного диалогического дискурса, так как при её разработке исследователи интересовали в первую очередь вопросы, которые «касаются внутреннего устройства устного дискурса, и, соответственно, предлагаемая транскрипция ориентирована на то, чтобы отразить членение дискурса на значимые и взаимосвязанные единицы» [10, с. 30]. Авторский коллектив также утверждает, что «дискурсивная транскрипция должна в максимальной степени отражать процессы порождения дискурса, как они разворачиваются в когнитивной системе говорящего. Для нас дискурсивная транскрипция – это модель речепорождения, реконструкция когнитивных процессов» [Там же, с. 31]. Как видим, при разработке данной системы аннотирования центральной фигурой признается говорящий, однако устный дискурс, описываемый в «Рассказах о сновидениях», относится к монологическому типу, который, в частности, принципиально отличается от диалогического количеством говорящих и – как следствие – возникновением проблемы соотношения их речевых партий. Таким образом, в рамках системы аннотации, разработанной к 2009 году, исследователи не имели возможности учесть в принципах аннотации специфику диалогического взаимодействия, вследствие чего не было предложено вариантов фиксации наложений, перебивов, случаев совместного построения реплик и прочих явлений, отмеченных ещё коллективом Е. А. Земской.

Исследовательская группа под руководством А. А. Кибрика и В. И. Подлесской продолжала работу над изучением устной речи и к концу 2018 года представила «ресурс “Рассказы и разговоры о грушах”, включающий серию записей естественной коммуникации между несколькими участниками, в совокупности с вокальной и кинетической аннотацией этих записей» [15]. Вокальная аннотация устной речи строилась на принципах, разработанных для устного монологического дискурса, так как речевая партия каждого участника диалога записывалась и аннотировалась отдельно. В январе 2019 принципы вокальной аннотации были дополнены, что обеспечило более адекватное представление дискурса диалогического типа в графической форме.

В работе Н. А. Коротаева «“Рассказы и разговоры о грушах”: принципы вокальной аннотации» предложено дополнительно учитывать две характеристики устного диалога, которые «напрямую касаются процесса вокального аннотирования и способов адекватного графического представления выполненной аннотации» [7, с. 21]: начиная с 2019 года в аннотации устного диалогического дискурса предлагается учитывать особый статус пограничных абсолютных пауз и наложений и вводится понятие партитурной нотации: «...для наглядного отображения описанных выше свойств и явлений в проекте наряду с индивидуальными вокальными транскриптами используется единая партитурная нотация. Партитурная нотация автоматически формируется в среде Excel» [Там же, с. 22].

Параллельно с этим с 2018 года на кафедре русского языка и общего языкознания Иркутского государственного университета началась работа по адаптации существующей системы аннотирования вокального канала для фиксации аудиозаписей глубинных интервью, полученных в результате экспедиции в Баргузинский район Прибайкалья.

Одними из первых явлений, специфических для устного диалогического дискурса, которые были зафиксированы в данной аннотации, стали наложения и паузы внутри речевой партии отдельного говорящего. Результаты первого опыта аннотирования были представлены в таблице MS Word и выглядели следующим образом (см. Таблицу 1).

Таблица 1. Предварительный вариант аннотирования

первый участник		второй участник		третий участник	
0,40-1,95 2,04-2,88	Давайте познакомимся <i>сначала</i> . {laugh}	1,36-2,44	Давайте познакомимся,	пауза 3,76 с	
пауза 0,88 с		2,44-3,76	Людмила Ниловна, я Марина.		
незаполненная пауза 0,26 с					
4,02-4,85	А вас как <звуют/звать>?	пауза 1,85 с		4,38-5,06	А я Ольга.
5,29-5,65	Ольга			пауза 0,71 с	
	звонит телефон	5,77-6,27	Ага.	5,77-6,27	Ага.
6,37-6,95	Хорошо.	пауза 0,72 с		пауза 0,72 с	

В Таблице 1 столбцы и строки позволяли отразить, какие именно реплики накладываются друг на друга. Сочетание полужирный + курсив позволяло обозначить более точные границы наложений, речь каждого участника была сегментирована, в аннотации фиксировались временные границы выделенных сегментов, для каждого участника размечались абсолютные паузы (периоды молчания).

Параллельное выделение несколькими исследовательскими группами одних и тех же специфических черт диалогического дискурса подтверждает необходимость их обязательной фиксации при аннотировании устных материалов такого типа.

После знакомства с результатами работы московской исследовательской группы принято решение по использованию предложенной Н. А. Коротаевым аннотации с некоторыми корректировками и дополнениями. Мы руководствовались принципом единообразия при описании и фиксации устного диалогического дискурса, который представляется очень важным, поскольку, с одной стороны, если в работах разных исследователей используется единообразное аннотирование материала, то это позволяет применять сходные методы анализа к материалам, полученным различными путями (эксперимент, интервью и т.п.), следовательно, обеспечивает возможность сопоставления имеющихся данных, а с другой – может впоследствии способствовать созданию подробно размеченного корпуса устных диалогических текстов разного типа. Помимо этого, в связи с высокой трудоемкостью работы, относящейся к аннотированию устного диалогического дискурса, и малым количеством тщательно обработанного материала, такое решение может оказаться полезным и для исследователей устной речи, по тем или иным причинам не имеющих возможности аннотировать записи диалогической речи подобным образом. Вследствие этого аннотации нашего материала приобрели следующий вид (см. Таблицу 2).

Таблица 2. Промежуточный вариант аннотирования

Line	t начала ЭДЕ	t конца ЭДЕ	ВИ	МБ	ОЛ	Комментарии	
169	02:37.58	02:38.88	Там же /потаск привязываешь,			произносит привязыв[а]шь	
170	02:39.59			/Чё привязываешь?			
171	02:40.15	02:40.36	хорошая це= == (э 0.34)				
172		02:41.14					
173	02:42.14		хороший \тросик (0.4) и потаск (0.2) специально,				
174	02:42.22	02:42.50				Ага.	
175		02:43.62					

В данном варианте аннотирования осуществлено разделение на ЭДЕ, выполнена их временная привязка, указаны просодические акценты и специфика движения тона в рамках ЭДЕ, зафиксированы наложения, отмечены индивидуальные произносительные характеристики носителя диалекта. Помимо этого, принято решение сохранить полужирное начертание для более точной фиксации границ наложений (вплоть до звуков), так как вариант аннотации, предложенный Н. А. Коротаевым, не позволяет отразить чёткие границы наложений.

Дальнейшая работа с материалом привела к тому, что мы обнаружили существенный недостаток аннотации вокального канала, предложенной Н. А. Коротаевым для дискурса диалогического типа. В ней оказались не учтены варианты графической передачи маркеров обратной связи. Так, в тексте статьи упоминается только о двух из них: «...используется конвенциализованная запись для сигналов подтверждения: угу и ага» [Там же, с. 5].

Как можно увидеть, в промежуточном варианте аннотации мы использовали те же варианты для графической передачи маркеров обратной связи, однако чем больше количество материала было проаннотировано, тем больше мы убеждались в необходимости разработки иного подхода к графическому оформлению плана выражения единиц такого типа. Во-первых, очень быстро стало понятно, что всё разнообразие устных планов выражения маркеров обратной связи просто невозможно свести к этим двум, во-вторых, как отмечала Е. А. Земская, маркеры обратной связи относятся к явлениям неканонической фонетики, это означает, что, например, при вокализации «утвердительно-междометия, произносимого с открытым ртом» [11, с. 26],

которое во всех исследованиях устной речи записывается как *ага*, между гласными никогда не произносится звук [г], канонически обозначаемый буквой *г*. Звук, который на самом деле произносится между гласными, описывается как «редкий в русском языке звонкий задненёбный фрикативный согласный, который может произноситься в некоторых словах (*господи, бога, ага, бухгалтер*), а также получается в результате озвончения [х] перед последующим шумным звонким согласным (*их бы*)» [6, с. 73]. В данном случае речь идет о *г* фрикативном – [ɣ]. В случаях с озвончением (*их бы, ах да, эх бы*) этот звук реализуется только благодаря ассимиляции, то есть функционирует в качестве варианта фонемы <х>. В словах *господи, бога* «возможность» реализации варианта [ɣ] можно было бы объяснить (в рамках сложившейся традиции) влиянием южнорусского диалекта, носителем которого мог бы быть говорящий, однако произношение этих слов с фрикативным [ɣ] регулярно встречается и в речи коренных сибиряков, вследствие чего такое объяснение кажется спорным, однако не возникает сомнений в том, что на данный момент существуют как варианты произнесения указанных слов со звуком [г], так и со звуком [ɣ]. Совсем иначе обстоит дело с маркерами обратной связи *ага* и *ого*. В наших аудиозаписях вышеупомянутые маркеры обратной связи произносятся исключительно с [ɣ], не встретилось ни одного случая, в котором бы [ɣ] был заменен на [г]. Помимо этого, можно предположить, что замена звука [ɣ] на [г] может повлечь за собой изменения в прагматической функции данного знака, однако это умозаключение основано на интроспекции, поэтому требует дополнительной проверки. Если же говорить об «утвердительном междометии, произносимом с закрытым ртом» [11, с. 26], которое в исследованиях устной речи записывается как *угу* или *мгм*, то первый вариант записи абсолютно не коррелирует со звуковым планом выражения, так как в этом маркере обратной связи оказывается невозможным обнаружить ни звук [у], ни звук [г], для этого нет никаких артикуляционных предпосылок, а если говорить о втором варианте, то, как кажется, использованию такого графического облика препятствует традиция употребления данного буквосочетания в художественной литературе: *Я даже готов дать вам удовлетворение, но только после победы. – Мгм, – неопределенно откликнулся Николай Степанович. – Скажите, милейший, а ваша затея с... э-э... ураном – на что нацелена?* (Андрей Лазарчук, Михаил Успенский. Посмотри в глаза чудовищ (1996)).

В приведенном примере буквосочетание *мгм* призвано маркировать задумчивость одного из собеседников. Довольно сложно сказать, какой именно вариант озвучивания предполагал автор этого текста, однако – теоретически – в данном контексте с помощью этого буквосочетания может быть передан как маркер обратной связи, о котором шла речь выше, так и нечто более близкое к *гм*, которое также принято использовать в текстах художественной литературы для обозначения задумчивости.

Такое положение вещей ставит под сомнение удачность выбора для фиксации в аннотации графического варианта *мгм*, который имеет оттенок значения задумчивости, так как аннотирование устного дискурса предполагает в том числе максимально точную фиксацию специфических явлений устной речи, а реализующиеся с высокой степенью частотности явления «неканонической фонетики», несомненно, стоит отнести к таким, которые составляют специфику русской устной речи.

Если сравнить звуковой состав маркеров обратной связи, которые принято записывать как *ага* и *угу*, то, помимо прочего, нет уверенности в том, что в них произносится один и тот же согласный звук. В международном фонетическом алфавите есть несколько звуков, способных выступать в качестве символического обозначения согласного в приведённых выше маркерах обратной связи, например, звуки [ɣ], [ʝ], [ɦ]. В частности, для графической фиксации маркера обратной связи, который часто обозначается как «угу», мы используем вариант *мицм*, так как полагаем, что на его артикулирование в значительной степени влияет позиция между сонорными, приближая его к аппроксиманту, а для записи маркера обратной связи, обычно обозначаемым как «ага», – вариант *ага*.

Стоит отметить, что первичную расшифровку некоторых фрагментов дискурса осуществляли студенты и, используя конвенционализованную запись для сигналов подтверждения, они очень часто ошибались при попытке соотнесения графического и звукового планов выражения маркеров обратной связи. Так, одни из них – независимо от того, с закрытым или открытым ртом был произнесен маркер обратной связи, – почти во всех случаях избирали вариант *ага*, а другие – *угу*. Это наблюдение также послужило в пользу решения отступить от конвенционального графического облика маркеров обратной связи. Заметим, что после того как был предложен альтернативный графический облик, дифференциация студентами в потоке речи этих маркеров обратной связи значительно улучшилась.

Нами предложен список обозначений (см. Таблицу 3), принятых в данном исследовании, для графической фиксации всех встретившихся в нашем материале маркеров обратной связи.

Таблица 3. Список символов, принятых в аннотации для обозначения маркеров обратной связи

1.	∧аа.	синонимичен <i>вот оно как / вот как</i>
2.	∧мм.	синонимичен <i>вот оно как / вот как</i>
3.	-а!	синонимичен <i>понял!</i>
4.	/а?	синонимичен <i>что?</i>
5.	-ага.	синонимичен <i>да</i>
6.	-да.	
7.	/да?	
8.	-да-да.	
9.	-да-да-да.	
10.	/как?	
11.	∖конечно.	

12.	/м?	синонимичен <i>что?</i>
13.	–мцм.	синонимичен <i>да</i>
14.	–мцм-мцм.	синонимичен <i>да-да</i>
15.	–м?м.	синонимичен <i>нет</i>
16.	–не.	синонимичен <i>нет</i>
17.	/не?	синонимичен <i>нет?</i>
18.	^не?а.	синонимичен <i>нет</i>
19.	–нет.	
20.	/нет?	
21.	/ничё себе!	
22.	–но.	синонимичен <i>да</i>
23.	–но-но-но.	синонимичен <i>да-да-да</i>
24.	\обязательно.	синонимичен <i>конечно</i>
25.	\понятно-понятно.	
26.	–/та-ак!	
27.	/что?	
28.	ясно.	

Как видим, все маркеры обратной связи, при произнесении которых регулярно реализуются явления неканонической фонетики, зафиксированы при помощи символов МФА (Международного фонетического алфавита), остальные же маркеры обратной связи переданы в соответствии с требованиями орфографии русского языка. Знаки слева от маркера обратной связи призваны обозначить характер движения тона (восходящий, ровный, нисходящий) при произнесении данных единиц, а знаки справа – их иллокутивно-фазовое значение.

Конечный вариант аннотирования можно увидеть в Таблице 4.

Таблица 4. Итоговый вариант аннотирования

Line	t начала ЭДЕ	t конца ЭДЕ	ЛС	МБ	НВ	Комментарии	
381	05:00.84						
382	05:00.98	05:01.38		–Мцм.	вот /пришёл вот,		
384		05:02.87					
385	05:03.36	05:04.39			с /армии внук.		
386	05:24.54		\Вот.				
387	05:04.70	05:04.93					
388		05:05.12		–Ага.			
389	05:05.52		И ни= /никуда не можем его устроить.				
390	05:07.10	05:07.95			Ну \краса-авец.		
391		05:08.27					

Заключение

Проведенное исследование позволяет прийти к следующим выводам. В результате анализа научной литературы, посвященной изучению устной речи, выявлен ряд лакун, существующих в теоретическом описании таких явлений, как обратная связь, наложения, перебивы, совместное построение реплик дискурса диалогического типа, паузы (как пограничные, так и паузы внутри речевой партии отдельного говорящего). Можно констатировать отсутствие как общепринятой системы аннотирования устного дискурса, так и системы аннотирования устного диалогического дискурса, которую можно было бы без адаптации и дополнений использовать при конвертации устного диалогического дискурса в письменную форму, осуществляя фиксацию всех явлений, релевантных для дискурса данного типа. Отсутствие единого подхода к описанию вышеперечисленных феноменов неизбежно влияет на прикладной аспект аннотирования.

В ходе решения второй задачи разработан такой способ отображения материала, который позволяет фиксировать феномены, характерные для диалогической речи, предложен ряд условных обозначений, способных максимально точно отражать эти явления.

Перспективы дальнейшего исследования мы видим в теоретическом описании и классификации обнаруженных явлений, специфических для устной диалогической речи, а также в расширении имеющихся корпусов звучащей речи и представлении их в открытом доступе для создания сбалансированной выборки, способной отразить особенности бытования русского устного диалогического дискурса.

Финансирование | Funding

RU Публикация подготовлена в рамках поддержанной Иркутским государственным университетом научной темы 111-21-301 «Устный корпус как инструмент лингвистических и междисциплинарных исследований».

EN The publication is prepared within research project 111-21-301 “Oral Corpus as an Instrument of Linguistic and Interdisciplinary Studies” supported by Irkutsk State University.

Источники | References

1. Богданова-Бегларян Н. В., Блинова О. В., Мартыненко Г. Я., Шерстинова Т. Ю., Зайдес К. Д., Попова Т. И. Аннотирование прагматических маркеров в русском речевом корпусе: проблемы, поиски, решения и результаты // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегод. Междунар. конф. «Диалог» (г. Москва, 29 мая - 1 июня 2019 г.). М.: РГГУ, 2019. Вып. 18 (25). С. 72-86.
2. Борисова И. Н. Русский разговорный диалог. Структура и динамика. М.: URSS, 2007. 317 с.
3. Звуковой корпус как материал для анализа русской речи: в 2-х ч. / отв. ред. Н. В. Богданова-Бегларян. СПб.: Филол. фак-т СПбГУ, 2013. Ч. 1. Чтение. Пересказ. Описание. 421 с.
4. Звуковой корпус как материал для анализа русской речи: в 2-х ч. / отв. ред. Н. В. Богданова-Бегларян. СПб.: Филол. фак-т СПбГУ, 2014. Ч. 2. Теоретические и практические аспекты анализа: в 2-х т. Т. 1. О некоторых особенностях устной спонтанной речи разного типа. Звуковой корпус как материал для преподавания русского языка в иностранной аудитории. 396 с.
5. Кибрик А. А., Коротаев Н. А., Федорова О. В., Евдокимова А. А. Единая мультимедийная аннотация как инструмент анализа естественной коммуникации // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегод. Междунар. конф. «Диалог» (г. Москва, 29 мая - 1 июня 2019 г.). М.: РГГУ, 2019. Вып. 18 (25). С. 288-304.
6. Князев С. В., Пожарицкая С. К. Современный русский литературный язык: фонетика, орфоэпия, графика и орфография. М.: Академ. проект; Гаудеамус, 2011. 430 с.
7. Коротаев Н. А. «Рассказы и разговоры о грушах»: принципы вокальной аннотации [Электронный ресурс]. URL: https://www.multidiscourse.ru/data/ann/pears_vocal%20annotation.pdf (дата обращения: 29.06.2020).
8. Коротаев Н. А. «Рассказы о сновидениях»: от автономной транскрипции к многоуровневой дискурсивной разметке // Корпусная лингвистика - 2011: тр. Междунар. конф. СПб.: СПбГУ, 2011. С. 205-210.
9. Подлесская В. И., Коротаев Н. А., Мазурина С. И. Самоисправления говорящего в русском монологическом и диалогическом дискурсе: опыт корпусного исследования // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегод. Междунар. конф. «Диалог» (г. Москва, 29 мая - 1 июня 2019 г.). М.: РГГУ, 2019. Вып. 18 (25). С. 547-562.
10. Рассказы о сновидениях: корпусное исследование устного русского дискурса / под ред. А. А. Кибрика и В. И. Подлесской. М.: Языки славянских культур, 2009. 736 с.
11. Русская разговорная речь. Тексты / отв. ред. Е. А. Земская, Л. А. Капанадзе. М.: Наука, 1978. 306 с.
12. Русская разговорная речь. Фонетика. Морфология. Лексика. Жест / отв. ред. Е. А. Земская. М.: Наука, 1983. 239 с.
13. Янко Т. Е. Коммуникативные стратегии русской речи. М.: Яз. славян. культуры, 2001. 382 с.
14. Янко Т. Е. Наречие 'давно' по данным звучащего корпуса // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной Международной конф. «Диалог» (г. Москва, 17-20 июня 2020 г.). М.: РГГУ, 2020. Вып. 19 (26), основной том. С. 773-784.
15. <https://www.multidiscourse.ru> (дата обращения: 07.05.2021).

Информация об авторах | Author information**RU****Михалёва Ольга Леонидовна**¹, к. филол. н., доц.
Чекмез Ульяна Эдуардовна²^{1,2} Иркутский государственный университет**EN****Mikhaleva Olga Leonidovna**¹, PhD
Chekmez Ulyana Eduardovna²^{1,2} Irkutsk State University¹ pikutskaja@gmail.com, ² chekmezu@gmail.com**Информация о статье | About this article**

Дата поступления рукописи (received): 15.05.2021; опубликовано (published): 30.06.2021.

Ключевые слова (keywords): устная речь; диалогический дискурс; аннотирование устного дискурса; обратная связь; маркеры обратной связи; oral speech; dialogic discourse; transcribing oral discourse; feedback; feedback markers.