

и использовались автором при написании научных работ [1]. Исследование подобных материалов требует особенно внимательного подхода, сопоставления с данными других источников, применения сравнительного анализа, но все же проведение такой работы представляется важным направлением в процессе изучения событий военных лет.

В заключение хочется отметить, что в рамках данной статьи нет возможности показать и даже просто перечислить весь спектр архивных материалов ТЦДНИ, содержащих сведения о повседневности военных лет. Многолетний опыт работы с материалами архива позволяет уверенно предполагать, что информационный потенциал фондов послужит надежной опорой для будущих исследований по различным вопросам, связанным с тематикой повседневной жизни советских граждан, переживших события Великой Отечественной войны.

Литература

1. *Болокина Л.А.* Жизнь населения оккупированных районов Калининской области в годы Великой Отечественной войны. Тверь: ТвГТУ, 2016. 160 с.
2. *Болокина Л.А.* Нарративные источники о Великой Отечественной войне в фондах Тверского центра документации новейшей истории // *Война и оружие. Новые исследования и материалы.* СПб.: ВИМАИВиВС, 2016. С. 282–293.
3. *Болокина Л.А.* Работа с беспризорными и безнадзорными детьми в Калининской области в 1942–1945 годах как социальная практика военного времени // *Вестник Оренбургского государственного педагогического университета.* 2018. № 2(26). С. 91–104.
4. ТЦДНИ. Ф. 83. Оп. 1. Д. 163.
5. ТЦДНИ. Ф. 147. Оп. 3. Д. 1148.
6. ТЦДНИ. Ф. 147. Оп. 3. Д. 1152.
7. ТЦДНИ. Ф. 147. Оп. 3. Д. 2197.

ПРИНЦИПЫ ТРАНСКРИБАЦИИ ШАХТЕРСКИХ ИНТЕРВЬЮ ДЛЯ АИС «АНТРОПОЛОГИЯ ТРУДА»

Логунова Н.В.,

*кандидат филологических наук, Национальный исследовательский
университет «Высшая школа экономики», г. Пермь,
e-mail: logunovan@yandex.ru,*

Пантелеева Л.М.,

кандидат филологических наук, доцент департамента иностранных языков, Национальный исследовательский университет «Высшая школа экономики», г. Пермь, e-mail: liliya_pant@mail.ru

Аннотация. В статье изложены принципы расшифровки материалов для наполнения АИС «Антропология труда». Система представляет собой ресурс для широкой пользовательской аудитории, аккумулирующий информацию по разным профессиям и презентующий ее в виде полных текстов источников, тематических отрывков, контекстуального кружения слов. Один из ее подкорпусов представлен собранием интервью с ветеранами горной отрасли Сахалина.

Принципы расшифровки интервью описываются как с точки зрения технического осуществления транскрибации, так и в отношении принципов сохранения информативности материалов, с тем чтобы они могли удовлетворять задачам и целям разноаспектных научных исследований. Данные, подготовленные таким образом для публикации в АИС, позволят специалистам разных научных сфер и направлений использовать их для решения соответствующих исследовательских задач.

Ключевые слова: антропология труда, антропологический корпус, автоматизированная информационная система, шахтеры, транскрибация.

PRINCIPLES OF TRANSCRIPTION OF MINER INTERVIEWS FOR AIS «ANTHROPOLOGY OF LABOR»

Logunova N.V.,

Candidate of Philological Sciences, HSE University, Perm, e-mail: logunovan@yandex.ru,

Panteleeva L.M.,

Candidate of Philological Sciences, Associate Professor of the Department of Foreign Languages, HSE University, Perm, e-mail: liliya_pant@mail.ru

Abstract. The article outlines the principles of deciphering materials for filling the AIS «Anthropology of Labor». The system is a resource for a wide user audience, accumulating information on various professions and presenting it in the form of full text sources, thematic passages, and contextual circles of words. One of its sub-corps is represented by a collection of interviews with veterans of the Sakhalin mining industry.

The principles of interview transcription are described both from the point of view of the technical implementation of transcription, and in relation to the

principles of preserving the information content of materials, so that they can satisfy the tasks and goals of multi-aspect scientific research. Data prepared in this way for publication in the AIS will allow specialists from various scientific fields and areas to use them to solve relevant research problems.

Keywords: *anthropology of labor, anthropological corpus, automated information system, miners, transcription.*

Введение. В рамках исследований трудовой повседневности человека цифровой корпус – явление новое и малоописанное. Приступая к созданию автоматизированной информационной системы «Антропология труда», мы стараемся осмыслить общую структуру такого собрания на примере одного из ее подкорпусов, материалами для которого послужили интервью с ветеранами горной отрасли Сахалина.

«Антропология труда» – это ресурс для широкой пользовательской аудитории, аккумулирующий информацию по разным профессиям и презентующий ее в виде полных текстов источников, тематических отрывков, контекстуального кружения слов. Главная задача такого ресурса – обеспечение пользователя данными по профессиональной культуре, отраженными в разных типах источников: устных и письменных, официальных и неофициальных, авторских и фольклорных и т.д.

Подкорпус шахтерских нарратив в настоящий момент представляет собой расшифрованные аудиофайлы, отобранные из всего множества интервью по принципу лингвокультурной ценности текста. Аудиозаписи были сделаны в рамках экспедиции «Социокультурная антропология шахтерского трудового быта в СССР» осенью 2022 года.

При расшифровке аудиозаписей нами была выработана система принципов, обеспечивающих сохранение разносторонней информативности материалов. Их описанию и посвящена настоящая статья.

Принципы расшифровки интервью для публикации в корпусе. Процесс транскрибации требует пояснений как с точки зрения его технического осуществления, так и в отношении принципов сохранения информативности материалов, с тем чтобы они могли удовлетворять задачам и целям разноаспектных научных исследований.

1. Технические принципы транскрибации.

На начальном этапе составления корпуса применяется ручная расшифровка, большинство принципов которой в дальнейшем могут быть перенесены в автоматизированные системы транскрибации.

1.1. Текст интервью должен предваряться своеобразным паспортом: ФИО респондента (полностью), год рождения, место работы, должность, место интервью, дата интервью, интервьюер, расшифровщик.

1.2. Текст интервью отделяется от «паспорта» пропущенной строкой. Каждый вопрос и каждый ответ печатаются с нового абзаца. Начало фразы дается с заглавной буквы. Конец фразы обозначается знаком « // ». Точка и другие знаки препинания не ставятся. В случае необходимости ставится знак «?». Микропаузы в тексте обозначаются знаком « / ».

1.3. Из других знаков препинания, помимо вопросительного знака, следует использовать кавычки в таких случаях, как:

а) при обозначении названий книг, кинофильмов, цитируемых строк из песен и т.п., например: *И 1: А песня любимая какая была / про шахту? // Р: Ну как у всех / «Там на шахте угольной» (смеется)* (1) и под.;

б) для обозначения реплик участников диалога. При расшифровке и письменной фиксации пересказанного респондентом диалога нужно заключать в кавычки реплики его участников. Это продиктовано тем, что респондент, выделяя чужую прямую речь интонационно, не всегда называет субъекта этой прямой речи. В аудиофайле эти реплики не сливаются, а в транскрипте, если их не выделить кавычками, слова одного лица могут быть приписаны другому: *Он ко мне подходит // «Я тебе сигнал куда давал?» // Я говорю / «На тебя» // «Что? // Ну-ка повтори мне» // Я говорю / «На тебя» // «Круговое движение я тебе даю» // Я говорю / «Ой Боря-а / как я неправильно сделал» // (1) и под.*

1.4. В ряде случаев фиксируется ударение в словах. Акцентологический вариант, произнесенный респондентом, нужно указывать, если:

а) вариант отступает от литературной нормы, например: *три человека чисто вывозят **дббычу** / уголь / вагонетками / вот // (1) и под.*

б) нормативное ударение помогает разграничить омофоны или грамматические варианты, например: *Ду́хи // Сейчас эти ду́хи там / они уже... // (2) (ср.: духи); тот уголь / который / из бункера **насыпали** // (1) и под.;*

в) респондент употребляет произносительные варианты, свойственные, вероятно, данной профессиональной речевой среде (это могут быть как акцентологические, так и лексические профессионализмы), например: *надо послать туда этих / **путевых** / вот // А в смену тоже ходили дежурные **путевые** / там два или три человека // (1); поезда ходили / **мотриса** ходила // <...> **мотриса** / это в народе так / это просто дизель-поезд // (1) и под.;*

г) респондент употребляет регионально специфичные обозначения – названия населенных пунктов, предприятий и т.п., или имена и фамилии нестандартных моделей, например: *В Томарі я слесарем устроился учеником / (3); Это наверно у нас главные шахтеры Ломáга и Алексеев // (2) и под.*

1.5. Аббревиатуры даются в тексте в том буквенно-звуковом варианте, который произнесен респондентом, например: *Одна половина была / значит / АБЭКа / административно-бытовой комплекс / (1) и под.*

1.6. Слова недоговоренные печатаются с так называемым «висячим дефисом» и многоточием, например: *У нее два принци-... // Она [гово]рит / два принципа должно быть / в семье шахтеров // (4) и под.*

1.7. Слова с утратой в середине нескольких звуков расшифровываются и даются в тексте в полном варианте, а утраченные в произношении звуки вставляются в квадратных скобках [], например: *она / ну наверно / была сантиметров ше[сть]де[сят] да / в высоту / (5); ну комбайн все-т[а]ки проходческий / (5) и под.*

1.8. Слова литературного языка, произнесенные респондентом в ошибочном варианте, помечаются знаком « * », чтобы они не воспринимались как опечатка, например: *какой-то полимер / такой да / был / вот / типа дермантина* что-то // (5) (норм.: дерматин); Специальные макишейдеры* ходили в шахту / (1) (норм.: маркшейдер) и под.*

При этом просторечные слова и формы, профессионализмы, жаргонизмы, вулгаризмы и обценная лексика никакими знаками не помечаются и приводятся в том виде, в каком их употребил респондент, например: *Вот там мы обучались // (1); Потому / пространство сжатое / вот // (1); Да они не хочут слушать // (3); ну / так называлось кабыргá / или мехлопата // (5); Он стоит / хайка у него отвисла / (3); Чё-то произошло непонятно что блядь // (4); Я / ни хера себе / подхожу / ему чуть фляжку в рот не засунул // (4) и под.*

Удвоенным знаком «**» отмечаются слова и фразы, которые респондентом не поясняются, и поэтому транскрибатор вынужден их комментировать в примечаниях, например: *ШИ** – шахтный интерферометр; ВэГэСэЧэ** – военизированная горноспасательная часть.*

Подобным же знаком помечаются явные оговорки респондента (типа *Могло бы ...случиться неповторимое*** Комментарий транскрибатора: *непоправимое (видимо, оговорка)*), а также грамматические фразы, которые в транскрипте могут быть восприняты как опечатка транскрибатора (*А чё мы про вас разговаривать**?* Комментарий транскрибатора: *<...> нам про вас разговаривать (устная конструкция, перестроенная по ходу речи)*).

1.9. Проявления невербальной коммуникации (смех, вздохи, жесты и т.п.) фиксируются с помощью ремарок в скобках, например: Р: *(смеется) Не-е-т / это сильно грубо (смеется) // (1)*; Р: *О-ой (вздыхает) / эти блокировки / это знаете / это постоянно // (2)*; Я вот здесь вот *(сопровождает объяснение жестами) / тут поворот / а вот здесь вот вагон // (1)* и под.

1.10. Посторонние шумы отображаются в тексте только в том случае, если участники интервью вербально реагируют на них, например: *Голос некоего мужчины, не И¹ и не Р (разговор, видимо, о появившемся человеке в форме): Это что за сопля такая? Прокурорская какая-то // И: Какая-то форма интересная // Р: Прокуроры / прокуратура // Прокурорская // он вишь весь зализанный // И: Важный (смеется) // Р: Зализанный / и прическа / и всё // (С издѣвкой) Я в прокуратуре работаю // (Слышен шум проезжающей машины. Возвращаются к обсуждению этой машины) Р: А чего она черная / поехала? // Обыкновенная черная машина // (4).*

1.11. Все фрагменты аудиозаписи, которые транскрибатору не удалось расслышать и вследствие этого расшифровать, отражаются в тексте с пометой «неразб.» и обязательно сопровождаются таймкодом. (При этом транскрибатор может дать свой предположительный вариант расшифровки неясно слышимого фрагмента, выделив его шрифтом красного цвета.) Например: *сразу получаю бумагу от этого (неразб. 1.21.20) / от заведующего значит этого / факультета / подготовительного / (неразб. 1.21.48) (2)* и под.

2. Принципы сохранения информативности материалов.

Если же говорить о принципах адекватной передачи разных типов информации, содержащейся в аудиозаписи, то именно этот аспект процесса транскрибации требует наибольшего внимания, поскольку речь идет о материалах, на основе которых могут производиться разноаспектные научные исследования. Задача транскрибатора в таком случае – максимально сохранить все возможные блоки разноплановой информации, которую можно извлечь из вербальных и невербальных составляющих зафиксированного аудиозаписью дискурса.

2.1. Принцип соблюдения фактологической точности. Для адекватного отражения незнакомых онимов, терминов, профессионализмов и т.п. необходимо привлекать различную справочную информацию. Транскрибатор, что совершенно понятно, может не знать терминов и профессионализмов, связанных с определенной профессиональной средой, и потому сомневаться в правильности их расшифровки. Относи-

тельно шахтерского языка проблема состоит в том, что пока не существует словаря подобной лексики – авторитетного и достоверного с точки зрения научно обоснованного подхода к сбору и интерпретации материала. Крайне мало и специальных исследований языка шахтеров. К таковым, пожалуй, можно отнести лишь диссертацию на соискание ученой степени кандидата филологических наук, выполненную Л.В. Прибытовой [1]. Очевидно, что описание профессионального языка этой сферы деятельности представляет собой явную лауну как в исследовательской, так и в лексикографической практике современных российских лингвистов.

Если же говорить не о решении исследовательских задач, а о конкретных случаях в процессе транскрибации аудиофайла – необходимости опознавания слова, которое неизвестно транскрибатору, то в этих целях можно воспользоваться материалами различных сайтов в интернете, например таких, как: (6), (7), (8), (9) и под. На них представлены словники профессиональной шахтерской лексики, в полноте и адекватности описания которой есть определенные основания сомневаться. Нужно иметь в виду, что в семантизации некоторых профессионализмов или жаргонизмов в этих словниках могут быть расхождения, объясняемые как различиями в региональных вариантах шахтерского языка, так и, возможно, разностью подходов авторов таких материалов к сбору и интерпретации профессиональной лексики. Но тем не менее, транскрибатору подобные интернет-источники помогут расшифровать незнакомое слово или словосочетание.

Рассмотрим пример 1: *ушел на Угледорскую **грозом** / вот // (5)*. Уточняем, есть ли такое обозначение – **гроз** – в шахтерском профессиональном языке. Википедия дает толкование: «Горнорабочий очистного забоя (ГРОЗ) – рабочая специальность в угольной и уранодобывающей промышленности. Одна из двух основных шахтерских профессий наряду со специальностью проходчика» (10).

Пример 2: ***верхняк** может быть / э-э / металлическим либо деревянным // (5)*. Объяснение профессионального жаргонизма **верхняк** находим в «Свободной шахтерской энциклопедии»: «**ВЕРХНЯК** – верхний элемент крепи (обычно арочной или кольцевой). Он же **ОГНИВО**» (8).

Таким образом, транскрибаторам можно пользоваться (с определенными оговорками) результатами так называемой гражданской лексикографии, пока квалифицированные специалисты не составили словари профессиональной шахтерской лексики.

Затруднения у транскрибатора в опознавании слова могут вызывать не только профессионализмы, но и достаточно редкие в употреблении жаргонные и просторечные слова. Так, например, предположительно расшифрованное транскрибатором и неизвестное ему слово **хайка** верифицировалось с помощью «Словаря русского арго» В.С. Елистратова: *ХАЙКА 1, -и, ж. Рот. От общеупотр. прост. «хаять», ХАЙ [2, с. 508]. Семантизация этой лексемы в словаре соответствует смыслу контекста: Он стоит / хайка у него отвисла / (3), и убеждает транскрибатора в правильности расшифрованного им слова.*

Отдельной проблемой для транскрибатора является расшифровка фамилий или незнакомых ему наименований различных объектов, допустим, прежних японских названий шахт. Очевидно, что такую информацию не всегда возможно уточнить. В этом случае в соответствующем месте транскрипта, выделенном красным шрифтом, проставляется тайм-код, с тем чтобы интервьюер, у которого, возможно, есть и письменные записи к интервью, мог прослушать данный фрагмент аудиофайла и устранить лауну в тексте транскрипта: *Ну там вот у нас шахты (с 13.32 и далее) Ханькó (возможно: Ханькоу) / Хакачитáла (возможно: Хакачизáна) / потом наша шахта называлась Мицубúси / потом шахта Ударнинская назвалась Таньхэ (возможно: Тайхэй?) / это вот японское название // (2).*

2.2. Принцип лингвистической информативности. Его соблюдение предполагает передачу в транскрипте всех особенностей устной речи и коммуникативного поведения респондента. Это чрезвычайно важно, поскольку речь человека, особенно зафиксированная в разнообразных обстоятельствах дискурсивного потока, высвечивает языковую личность говорящего в совокупности всех его речевых компетенций, социобиопсихических и ментальных характеристик. Отражение в транскрипте полной информации о речевом поведении респондента позволит лингвистам воссоздать речевой портрет его языковой личности, а, допустим, социологам или культурологам исследовать уровень общей и профессиональной культуры и аксиологические приоритеты представителей разных профессиональных сообществ.

Исходя из сказанного следует сохранять и отражать в транскрипте:

а) особенности произношения, например: чё, ничё, ско[ль]ко, г[ово]рит и т.п.,

б) особенности постановки ударения, например: *чтобы / э-э / крутились конвейерá / (2); вот / и у шоферóв плечи / отнимаются // (3); Как у нас говорили свежачок / только **úз** леса // (5) и под.;*

в) слова-паразиты; звуки, заполняющие паузы; недоговоренные слова и фразы; самоперебивы и т.п., например: *И вот только свет электровоза / свет / э-э / эт самое / налобной лампы / и всё* // (1); *И потом / я посмеялся / потому что ...* // *А потом мне Серёдкин говорит* // (2); *А-а / кого-нибудь там / кто-то там кого-то там обсуждает / чё-то там / как / э-э / вчерашний день прошел / как это самое* // *Ну и женщины / это самое ...* // (2) и под.;

г) все отступления от нормативного литературного словоупотребления (просторечные, диалектные, обценные слова и выражения), а также профессиональную лексику, например:

– просторечные лексемы и грамматические формы: *Вот трикушку / неделю потаскаешь / домой приносишь / она вся белая* // (3);

– диалектные слова: *И сейчас вот я надысь его видел* // (3);

– обценные слова и выражения: *Они там матом / да ёб твою мать* // (2);

– профессионализмы: *короче заштыбовался конвейер* // (2) и под.

2.3. Принцип отражение особенностей коммуникативного поведения и прагматических установок респондента. К этому относится:

– фиксация (с помощью ремарок в скобках) особенностей коммуникативного поведения респондента, например, его эмоциональных реакций в виде смеха, вздохов, определенных жестов (если о них можно судить по речевому материалу): *Преподаватели перепугались / что за крик (смеется)* // (3); *Во-о-т / это / о-о-й (вздыхает) / раз[в]е вспомнишь / сколько лет* // (3), или жестикуляции, дополняющей и уточняющей какие-либо описания, например: И²: *А Загорск это где / туда дальше что ли?* // Р: *Вот получается сопка ... (показывает жестами)* // (11) и под.

– отражение особых способов произнесения отдельных слов (например, растягивание гласных звуков в слове или произнесение слова по слогам и под.), а также интонации (иронической, торжественной и т.п.), демонстрирующих то или иное оценочное отношение респондента к предмету речи: *А шахтеры отмечают праздник / о-очень / о-о-очень / плодотворно / ага / честно отмечают (смеются)* // (2); *(Передразнивает пьяного) Бе-бе-бе / ну вчера бухнули* // (4); *Зализанный / и прическа / и всё* // *(С издёвкой) Я в прокуратуре работаю* // (4) и под.;

– описание ситуаций, которые возникают по ходу интервью, если речевой материал показывает, что респондент на них реагирует, например: *(звонок телефона) Мой или ваши?* // *Мой?* // (4).

Заключение. Таким образом, если транскрибатор полно и точно отразит в расшифровке всю фактологию, все особенности устной спонтанной речи и коммуникативного поведения респондента, то это позволит специалистам разных научных сфер и направлений использовать данные материалы для решения соответствующих исследовательских задач (в частности, например, для изучения состояния общей, профессиональной и речевой культуры того или иного представителя данной социальной группы). Лингвисты, например, на основе анализа отраженных в транскрипте индивидуальностей речевого поведения респондентов (умения строить связный нарратив, находить адекватную форму для выражения собственных мыслей и эмоций и т.д.) смогут произвести лингвоперсонологическое портретирование как отдельной языковой личности, так и определенного профессионального сообщества.

Очевидно, что соблюдение перечисленных принципов транскрибации аудиофайлов сделает собранный материал многопланово информативным и потому ценным для специалистов разных гуманитарных наук – антропологов, социологов, лингвистов, историков материальной и духовной культуры и т.д.

Источники

1. Транскрипт интервью Еремина В.О., машиниста электровоза шахты Долинской, с. Быков, Сахалинская область.
2. Транскрипт интервью Болдырева В.А., электрослесаря шахты Ударновской, г. Шахтерск, Сахалинская область.
3. Транскрипт интервью Федосова В.Н., машиниста электровоза шахты Горнозаводской, г. Горнозаводск. Сахалинская область.
4. Транскрипт интервью Артамонова В.С., начальника участка шахты Углегорской, г. Шахтерск, Сахалинская область.
5. Транскрипт интервью Бородина В.А., мастера по вентиляции и технике безопасности шахты Углегорской, г. Южно-Сахалинск, Сахалинская область.
6. Словарь горных терминов и понятий // Яковлев Г. Никита Изотов. Сайт «ВикиЧтение». URL: <https://biography.wikireading.ru/232889> (дата обращения: 08.03.2024).
7. MiningWiki – свободная шахтерская энциклопедия. URL: https://miningwiki.ru/wiki/Заглавная_страница (дата обращения: 08.03.2024).
8. Шахтерский жаргон // Сайт «livejournal.com». URL: <https://coal-liza.livejournal.com/21835.html> (дата обращения: 08.03.2024).

9. Шахтерский слэнг // Сайт «dzen.ru». URL: <https://dzen.ru/a/YenuXw0hzAdqX7xz> (дата обращения: 08.03.2024).
10. Горнорабочий очистного забоя // Сайт «ru.wikipedia.org». URL: https://ru.wikipedia.org/wiki/Горнорабочий_очистного_забоя (дата обращения: 08.03.2024).
11. Транскрипт интервью Поманисточка А.И., мастера шахты Долинской, г. Долинск. Сахалинская область.

Литература

1. *Прибытова Л.В.* Профессиональный язык шахтеров Кемеровской области: Терминология и жаргон. Автореф. дисс. ... к. филол. н. Кемерово, 2005. 22 с.
2. *Елистратов В.С.* Словарь русского арго: Материалы 1980–1990 гг.: Около 9000 слов, 3000 идиоматических выражений. М.: Русские словари, 2000. 694 с.