

Фадеева Л. В.

ЭТАПЫ ПОДГОТОВКИ РЕЧЕВОГО МАТЕРИАЛА К ЛИНГВИСТИЧЕСКОМУ АНАЛИЗУ

Адрес статьи: www.gramota.net/materials/1/2007/3-3/102.html

Статья опубликована в авторской редакции и отражает точку зрения автора(ов) по рассматриваемому вопросу.

Источник

Альманах современной науки и образования

Тамбов: Грамота, 2007. № 3 (3): в 3-х ч. Ч. III. С. 237-238. ISSN 1993-5552.

Адрес журнала: www.gramota.net/editions/1.html

Содержание данного номера журнала: www.gramota.net/materials/1/2007/3-3/

© Издательство "Грамота"

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: www.gramota.net

Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: almanac@gramota.net

5. **Краснова М.** Искусство авангарда. Генезис изобразительных форм // Техническая эстетика и промышленный дизайн, № 4, 2006.
6. **URL:** <http://www.russiansnj.com/modules.php?op=modload&name=New&file=article&sid=8373&mode=thread&order=&thold=0>
7. **Роб-Грийе А.** Встреча. М., 2001.
8. **Кино.** Энциклопедический словарь. М., 1986.
9. **Бунтман Н.** Неизменное место встречи // Роб-Грийе А. Встреча. М., 2001.

ЭТАПЫ ПОДГОТОВКИ РЕЧЕВОГО МАТЕРИАЛА К ЛИНГВИСТИЧЕСКОМУ АНАЛИЗУ

Фадеева Л. В.

Астраханский государственный технический университет

Прежде чем приступить к непосредственному анализу устной речи, необходимо собрать и подготовить материал, определить его объём, возможности, степень детализации и связи, существующие в поле исследования. Современные компьютерные технологии многократно упрощают и ускоряют процедуры лингвистической обработки больших массивов устных текстов.

Исследование особенностей немецкой устной речи тесно связано с корпусной лингвистикой и опирается на речевой корпус. Сфера применения **корпусной лингвистики** обширна, т.к. ценные ресурсы, которыми располагают корпуса, позволяют абсолютно всем языковедам проводить не только дескриптивные, теоретические, но и прикладные лингвистические исследования. Для многих отраслей лингвистики корпуса являются ценными источниками: они используются для создания словарей, изучения языковых изменений и отклонений, понимания процесса языкового заимствования и т.д.

Для проведения анализа устной речи нередко используются уже имеющиеся корпуса устной речи, т.к. сбор данных является очень трудоёмким и дорогостоящим процессом. **Корпуса** представляют собой машинно-считываемое собрание документов естественной речи, которые могут быть приведены в единообразный вид, дополнять друг друга или контрастировать. Многие корпуса общедоступны как тома текстов в электронном виде, на аудиокассетах, дисках или в интернете, пользование некоторыми корпусами возможно только с разрешения исследовательских учреждений.

Самые известные разговорные корпуса - английский корпус London-Lund [Svartvik, Quirk 1980] и «Тексты устного немецкого литературного языка I-IV» [Texte gesprochenen deutscher Standardsprache I-IV, 1971]. Кроме того, есть корпуса, в которых собраны разговорные дискурсы определённых типов или свойственные отдельным группам говорящих, например, разговор-консультация, молодёжный дискурс или разговор на производстве, многие из них опубликованы в серии книг Phonai-Reihe тюбингенского издательства Niemeyer. Гамбургский компьютерный тезаурус (НСТТ - Der Hamburger computergestützte Transkriptionsthesaurus) наряду с транскриптами записей устной немецкой речи содержит также транскрипты на иностранных языках (турецком, французском и т.д.), составленные на основе транскрипционной программы SYNCWRITER, разработанной в Гамбурге. Корпуса Баварского Архива речевых сигналов в Мюнхене (BAS - Das Bayerische Archiv für Sprachsignale) делятся на: корпуса читаемого языка, корпуса спонтанной речи и корпуса акцентной / диалектной речи. Ульмский университет обладает банком текстов, в котором хранятся орфографические транскрипты записей психотерапевтических бесед.

Самым большим фондом устных разговорных корпусов в Германии располагает архив немецкого языка (Archiv für Gesprochenes Deutsch) мангеймского Института немецкого языка, часть его корпусов уже подверглась компьютерной обработке и доступна в интернете в виде печатных и звуковых банков данных. В мангеймском архиве находятся большие корпуса регионально окрашенной разговорной речи и устного литературного языка, звуковые записи и транскрипты вербальной интеракции в различных социальных и ситуативных контекстах. В настоящий момент в архиве находятся 39 корпусов устной речи, около 900 видеозаписей, 16300 аудиозаписей с общей продолжительностью звучания 4400 часов, а также около 6650 транскриптов.

Проблемой речевых корпусов является редукция первоначальных данных (реальных бытовых бесед). Любая техническая фиксация интеракции является лишь упрощённым вариантом реальности (Reduktion von Realität), её моделью, которая на несколько порядков беднее, чем сама действительность [Stachowiak 1973: 36].

Однократного наблюдения не достаточно для проведения анализа, чтобы преодолеть комплексность первичных данных, они многократно обрабатываются и подготавливаются к научному анализу. Таким образом, возникают вторичные данные, которые представляют собой редуцированные модели оригинала первичных данных. Первичные данные фиксируются для повторной аналитической выборки в различных формах письменных протоколов. Для анализа устной речи интерес представляют только те формы вторичных данных, которые записаны на аудио- или видеоплётку. Коммуникативные события, представленные в устной форме, отличаются комплексностью и ситуативностью, поэтому анализ, опирающийся лишь на использование аудио- и видеозаписей, всё же является недостаточно точным. Только письменная графическая фиксация делает возможным полноценный, всесторонний анализ. Из аудио и видеозаписей готовятся транскрипты, которые являются третичными данными, лишь затем возможен анализ. Из этих транскриптов

можно создать упрощённые отредактированные транскрипты или другие графические модели. Таким образом, сбор всех данных необходимых для анализа, проходит поэтапно и на каждом отрезке представляет собой специфическую редукцию первичных данных:

Первичные данные → *Вторичные данные* → *Третичные данные* → *Анализ*.

Транскрипция имеет особое значение для анализа устной диалогической речи. В ней полученные данные дальше редуцируются и модифицируются. Метод транскрипции позволяет более точно определить, какие аспекты коммуникации необходимо проанализировать. По сравнению с записями, качество которых зависит от выбора технических средств, и которые в свою очередь окончательно определяют характер материала, решения, принятые относительно транскрипционных систем возможно пересматривать.

К. Бринкер выделяет два совершенно разных метода составления транскрипции: «пошаговая запись текста» (Textnotation) и «пошаговая запись партитуры» (Partiturnotation). Оба метода могут быть представлены как простая система или как построчная система блоков (Zeilenblocksystem). При пошаговой записи текста слова говорящих записываются в отдельных блоках текста. Перед каждым таким текстовым блоком пишется сокращённое обозначение соответствующего коммуниканта [Brinker 2001: 40]. Такая пошаговая запись текста используется, например, в корпусах текстов Фрайбургского исследовательского центра устной речи. Пошаговая запись текста хорошо подходит для протяжённых речевых высказываний с небольшим количеством собеседников, и для исследований, в которых речь идёт не столько об интерактивном взаимодействии отдельных высказываний, сколько об их внутренней структуре. И, наоборот, если в разговоре принимает участие много собеседников, которые к тому же используют короткие высказывания и частую смену ролей и в которых достаточно часто выступает параллельное говорение, то к таким разговорам применимы системы партитур. Эти системы позволяют лучше изображать интерактивные структуры. В отличие от метода пошаговой записи, их сложнее читать, и, если высказывания длинные, то они достаточно сложны и громоздки в своём составлении.

В силу того, что транскрипция, даже если она ограничена только вербальным аспектом, является трудоёмким и требующим затрат времени методом, целесообразно транскрибировать только то, что необходимо для дальнейшего анализа. Если беседа рассматривается как целое, в котором выделяются структуры (например, общий процесс аргументации), то необходимо транскрибировать весь разговор. Когда такие транскрипты готовы, речь идёт о «**речевом корпусе**» (Gesprächskorpus).

С другой стороны в рамках исследования устной диалогической речи решается целый ряд вопросов, ограниченных узкими локальными событиями разговора. Необходимо лишь выделить отрывки из длинной интеракции, в которых происходит интересующее нас событие. Для такого вида анализа нужен «**корпус событий**» (Ereigniskorpus) - коротких отрывков, выбранных из достаточно обширного материала, чтобы затем детально их транскрибировать. Для этого предусмотрено создание **опорных протоколов** (Orientierungsprotokolle).

Опорные протоколы - это метод пошаговой записи, который ещё в нечёткой разговорно-бытовой форме описывает процесс интеракции. Опорные протоколы готовятся посредством точного прослушивания и просмотра записанного на аудио- или видеоплётку разговора. Они служат для того, чтобы получить первый, но уже достаточно точный обзор всего материала и на его основе выбрать интересные событийные отрывки.

Такие опорные протоколы позволяют выбрать в материале интересные отрывки, слушать и просматривать их ещё раз, а затем решить, транскрибировать их или нет. Если корпус событий составлен таким способом, то можно транскрибировать и анализировать отдельные отрывки.

При точном техническом и методическом выполнении эмпирического исследования в рамках анализа устной речи следует строго соблюдать эτικο-юридические нормы исследования. Болевой точкой как юридической, так и этической проблематики является, по сути, альтернатива открытой или скрытой записи диалогов, имеющая фундаментальное значение для развития теории анализа устной речи и проведения эмпирических исследований. Исследователю приходится делать выбор между оптимальным для развития теории и допустимым с эτικο-юридической точки зрения решением, поскольку то, что необходимо для проведения научных изысканий, как правило, проблематично или абсолютно неприемлемо с эτικο-юридической точки зрения.

Тем не менее, существует возможность делать записи в рамках анализа устной диалогической речи. С юридической точки зрения не вызывает проблем запись, на которую дано согласие участников беседы. Если же с точки зрения методики получение разрешения перед записью неблагоприятно и вредит аутентичности записи, то разрешение должно быть обязательно получено после записи.

Список использованной литературы

1. **Brinker K.** (2001): Linguistische Gesprächsanalyse: eine Einführung / von Klaus Brinker und Sven F.Sager. - 3., durchges. und erg. Aufl. - Berlin: Erich Schmidt.
2. **Stachowiak H.** (1973): Allgemeine Modelltheorie. Wien, New York.
3. **Svartvik J., Quirk R.** (1980): A corpus of English conversation. Lund.
4. **Texte gesprochener deutscher Standardsprache I-IV.** (1971) 4 Bde. München.