

СЕКЦИЯ “НОВЫЕ НАПРАВЛЕНИЯ В ИССЛЕДОВАНИЯХ ДИСКУРСА И ДИСКУРСИВНОЙ ТИПОЛОГИИ”

В.Е. Гольдин, О.Ю. Крючкова

ТЕМАТИЧЕСКАЯ РАЗМЕТКА И ТЕМАТИЧЕСКИЙ АНАЛИЗ ДИАЛЕКТНОГО ТЕКСТОВОГО КОРПУСА*

Саратовский государственный университет

Одним из важных параметров метаразметки (метаописания) в автоматически обрабатываемых текстовых корпусах является тематическая разметка. Этот параметр применяется в корпусной лингвистике, в частности, в Британском национальном корпусе и в Национальном корпусе русского языка [10]. Задача создания диалектного текстового корпуса также диктует необходимость выработки принципов тематической классификации диалектных текстов. При этом, с одной стороны, важно учитывать уже имеющийся в корпусной лингвистике опыт кодирования предметной области текста. С другой стороны, тематическая разметка в диалектном корпусе не может полностью совпадать с рубрикацией, используемой для метаописания письменных текстов, ввиду специфики устной речи в целом и диалектной ее разновидности в частности.

Тематическая рубрикация материала принята в диалектологических хрестоматиях для указания главных содержательных компонентов публикуемых фрагментов речи и выполняет в хрестоматиях роль заголовков. Так, в изданной Институтом русского языка им. В.В. Виноградова звучащей хрестоматии, посвященной южно-русскому наречию [9], указаны общие темы каждого из рассказов: “О жизни во время войны”, “О семье мужа”, “О конопле”, “Свадебный обряд”, “О том, как слух потеряла”, “О своей жизни”, “О пожаре” и др. Поскольку тематика речи на диалекте в подобных изданиях не является предметом научного анализа, то к их тематической руб-

* Работа выполнена при поддержке Российского фонда фундаментальных исследований (РФФИ), проект № 06-06-80428-а.

рикации не предъявляются строгие требования последовательности и системности. Специальному изучению тематическая и тесно связанная с ней жанровая организация диалектной речи подвергаются при создании диалектных текстовых корпусов и в работах, посвященных языковой личности носителя диалекта [1; 2; 3].

Текстовый диалектный корпус должен служить моделью традиционной сельской коммуникации на диалекте, а если он включает текстовые материалы одного говора, то моделью коммуникации в конкретных условиях данного речевого коллектива. В этом случае важно учесть и реальное тематическое разнообразие речи, и хотя бы примерное количественное соотношение различных тем и жанров в континууме сельского речевого общения. Применительно к диалекту решение этой задачи оказывается особенно сложным, так как диалектологи почти всегда получают магнитофонные записи диалектной речи в ходе включенного наблюдения и в значительной мере сами определяют выбор обсуждаемых тем (см., например, приведенный ниже короткий диалог об именах лиц в телесериалах).

Однако положение не безнадежно: диалектолог определяет главную, общую тему разговора, но мало влияет или почти не влияет на тематические отступления и тематические переключения, чрезвычайно характерные для речи носителей диалекта, а эти отступления и переключения содержат объективную информацию об интересах и общих тематических предпочтениях рассказчиков, о наиболее актуальных для респондентов сторонах жизни. Это обстоятельство требует своего учета при разработке принципов тематической разметки текстов диалектного корпуса. Тематическая разметка диалектного корпуса не может быть одноуровневой и одноплановой, она должна различать широкие общие темы и узкие частные, предлагаемые исследователями и вводимые самими респондентами, формулируемые собирателем и называемые носителем диалекта, основные темы и фоновые и т.д.

Поскольку диалектная речь характеризуется предельной степенью разговорности, то при разработке принципов тематической разметки материалов диалектного корпуса необходимо учитывать достижения исследователей разговорной речи (см., например: [11; 8; 4; 5] и др). Применительно к разговорной речи не только определена специфика ее тематической организации, но и создана ее жанрово-тематическая типология. Попытки экстраполяции этой типологии на диалектный материал выявляют не только общее, но и существенные различия. Они связаны, прежде всего, с тем, что в русской деревенской традиции то, что соответствует быту, производству, досугу и т.п. сторонам городской жизни, четко не разграничено в темпоральном, территориальном и социальном плане. Поэтому время и место обще-

ния и даже состав общающихся в традиционной сельской коммуникации тематически обычно менее значимы, чем в городской. Деревенскому общению, кроме того, свойствен в целом гораздо больший уровень общности знаний, на которых базируется взаимопонимание сельчан.

Диалектный текст в целом и отдельные его фрагменты принципиально политематичны, в них затрагиваются одновременно несколько предметных областей. Вместе с тем с точки зрения грубой (широкой) тематической рубрикации диалектные тексты достаточно однородны, что обусловлено функциональной ограниченностью коммуникации на диалекте бытовой сферой общения. Предметная специфика диалектного текста обуславливает целесообразность его двухуровневой тематической разметки – широкой и узкой. Широкую разметку имеет смысл максимально приблизить к тематическому кодированию, применяемому в поиске на массивах письменных текстов, что обеспечит сопоставимость различных корпусов и включенных в них текстов. Узкая разметка должна отражать тематическую структуру широкой предметной области, выявляя ее специфику. Узкая тематизация послужит также базой для лексико-семантических и когнитивных исследований диалектной речи.

Тематически размеченные корпуса диалектных текстов могут иметь различную структуру. Тематическое деление может быть положено в основу членения корпуса, как это делается в диалектном подкорпусе Национального корпуса русского языка. В этом случае снимается фактор информанта, и корпус делится на тематические блоки, включающие текстовые фрагменты, записанные от разных информантов [см.: 6, 48; 7]. Тематическое членение может носить вспомогательный характер. При этом основной единицей корпуса является единовременная запись диалектной речи, произведенная от одного или (реже) нескольких информантов. У того и другого подхода есть свои преимущества. Принцип собственно тематической рубрикации предпочтителен для исследований культурологической направленности. Сохранение целого текста, представляющего естественную речь носителя народной речевой культуры, дает больше возможностей для изучения особенностей диалектной коммуникации, ее когнитивно-дискурсивной специфики. Этот подход принят в диалектном текстовом корпусе, создаваемом на кафедре теории, истории языка и прикладной лингвистики Саратовского государственного университета.

Задачей данной статьи является апробация второго подхода к тематическому членению диалектной речи. В статье предпринята попытка тематического кодирования и тематического анализа значительного по объему диалектного текста (45 мин. звучания, 59 тыс. знаков в расшифровке аудиозаписи), записанного в 1999 г. в с. Бело-

горное Вольского района Саратовской области от уроженки этого села 1939 г.р. Запись представляет собой беседу, происходившую на улице при случайной встрече небольшой группы молодых диалектологов с сельской жительницей.¹

На уровне широкой тематической разметки в анализируемом тексте выделяются следующие классификационные рубрики: *частная жизнь, дом и домашнее хозяйство, религия, зрелища и развлечения, политика и общественная жизнь, производство*. Названные предметные области занимают в тексте неодинаковое место и с точки зрения объема, и с точки зрения их смыслового разнообразия, т.е. возможности членения на микротемы, и с точки зрения особенностей вербализации соответствующей тематики.

Наиболее объемными являются тематические блоки *частная жизнь, дом и домашнее хозяйство, религия*. Остальные темы не получают сколько-нибудь значительного развития или возникают не как самостоятельные, а попутно, в рамках других тематических фрагментов. Так, *производственная и общественно-политическая* темы возникают в рамках тематики *частная жизнь*, тема *зрелища и развлечения* затрагивается попутно в рамках тематического блока *домашнее хозяйство*: Эти несамостоятельно развертываемые темы отличаются краткостью (100-200 знаков) на фоне подробных повествований (объемом в 1-2 тыс. знаков), касающихся вопросов частной жизни, домашнего хозяйства и религии. Ср.: (1) *...а хорошего тут правильно не дождёмся ничего // вот дошли / всё развалили / весь колхоз Z / а теперь / вон говорят / и по радио и по телевизору / “восстанавливать колхозы расширять их” / каким Z их будешь /;* (2) *... вечером вот в семь / в восемь часов / как дань отдай // кино поглядишь в шесть часов и выходи убираться //*

Объем текстового пространства, занимаемого той или иной темой, не находится, однако, в безусловной и прямой корреляции с когнитивной актуальностью данной тематики. Роль различных предметных областей в ассоциативно-вербальной сети говорящего выявляется также с помощью узкой тематической разметки, конкретизирующей смысловую объем широких тематических блоков, и в результате анализа дискурсивной специфики данного тематического сегмента текста.

Узкая тематическая разметка обнаруживает различную степень детализации затрагиваемых в тексте предметных областей. Темы *производство, политика и общественная жизнь*, а также тема *зрелища и развлечения* отличаются содержательной ограниченностью. Произ-

¹ В дальнейшем слова диалектолога маркируются буквой Д., слова информанта – буквой И, нецензурные слова – буквой Z.

водственная и общественно-политическая тематика маркированы только лишь упоминанием о профессии, темой низкой пенсии и темой разрушения колхозов. Тема *зрелища и развлечения* ограничена разговором о телесериале. В противоположность этому наиболее объемные тематические блоки (*частная жизнь, дом и домашнее хозяйство, религия*) характеризуются значительным смысловым разнообразием, включая в свой состав целый ряд развернутых микротем. В рамках широкой темы *частная жизнь* выделяются микротемы *родители, дети, другие родственники, брак и семейная жизнь, односельчане*. Тематический блок *дом и домашнее хозяйство* включает подтемы *огород, домашние животные, строения, техника и орудия труда*. Тема *религия* подразделяется на микротемы *православные и иноверцы, религиозные праздники и обряды, религиозный этикет, места религиозных отправлений*.

Степень смысловой детализации предметной области, несомненно, указывает на ее когнитивную актуальность. Вместе с тем этот параметр так же, как и объем тематических фрагментов, не является достаточным для определения места предметной области в индивидуальном речевом сознании говорящего. Значимыми в этом отношении являются лингвистические особенности, характеризующие соответствующий тематический отрезок текста.

Так, предметная область *зрелища и развлечения*, представленная единственной узкой темой *телесериалы*, заметно отличается по характеру ее вербализации от того дискурсивного типа, который обнаруживается в рамках других тематических блоков и может быть описан как активная, развернутая, эмоциональная речь, избыточная многообразными ассоциативными связями и усилительными повторами. Ср. фрагмент, непосредственно предшествующий разговору о сериалах, в котором представлено типичное для информанта построение текста, склонность к ассоциативному развертыванию речи, к выходу за рамки ответа на вопрос, предложенный собеседником:

(3) Д: А как Вы телят называете? Ну как Вы их зовете?

И: Вот с малолетства если назовёшь Мишка / то и до старости его зови Мишка / как уж его уничтожишь // они понимают / а и говорю / щас / дочка / скотина / умнее людей // я так вот соблазнила / ей-богу // если человек вот / пыряй ему в глаз / ему / говорит / божья роса // а скотина вот она / видишь / она попросилась / она знает время / её напоил / она всё спокойна / она больше вот не просит // и всё // а человек / если он загуляет / ему давай и давай / он до тех пор как / у него / бросит // неделю / неделю / две / две / зарядит / порет // вот // “да хватит / что-ля / хватит” // “тут просит” / говорит / (#понижает голос#) вот есть / вот он тут / (#изображает хриплым, грубым голосом#) “чай вы говорит когда говорит чё захотите / рыбы захотите аль колбасы аль чё-ли захотите / вы / говорит / покупаете али делаете

чего сами / и едите // а мы курить пить хотим вот / у нас тоже говорит просит там / червяк” // вот и всё // это вот...

В пределах же тематического фрагмента **телесериалы** наблюдается совершенно иная речевая манера, для которой все перечисленные выше характеристики должны быть даны со знаком “минус”. Информант не проявляет речевой активности, ограничивается односложными ответами на вопросы собеседников, в ответных репликах сельской жительницы отсутствуют эмоционально-оценочные средства, затронутая тема не является базой для развития каких-либо ассоциаций. Ср.:

(4) Д.: А Вы сериалы смотрите, да?

И.: Глядим / да //

Д.: А нравятся?

И.: Кто?

Д.: Сериалы.

И.: Да //

Д.: А имена в сериалах нравятся? Как вот зовут там людей?

И.: Марсело // тот / как уж её ещё -то / ну...

Д.: Эдуардо.

И.: Эдуардо // Альберто там есть // эта... как её... Лаура / да // не знаю / Лаура-то чё она? Не родила ещ что ля? не знай // я вот два дня не глядела //

Не меняется данная речевая манера и после реплик диалектолога, побуждающих собеседницу к пересказу сюжета фильма. Тема **телесериалы** не получает текстового развития, несмотря на то, что затронута она самим информантом и лишь подхвачена диалектологами. Следует также заметить, что в продолжение длительной беседы к этой теме информант больше не возвращается, тогда как все другие темы, выделенные в анализируемом тексте на уровне широкой тематической разметки, неоднократно поднимаются в ходе беседы по инициативе самого информанта, хотя и обсуждаются с разной степенью подробности. Заметное дискурсивное отличие тематического фрагмента **телесериалы** в анализируемом тексте позволяет сделать вывод о незначительной роли данной тематики в ассоциативно-вербальной сети информанта.

Специфика дискурсивного выражения производственной и общественно-политической тематики носит другой характер. Она ограничивается такими признаками, как несамостоятельное развертывание, вхождение в состав тематического блока **частная жизнь**, смысловая ограниченность. Вместе с тем текстовые отрезки, вербализующие данную тематику, не отличаются по своим лингвистическим характеристикам от микротекстов, отражающих тематику основных предметных областей. Это наблюдение позволяет предположить, что

специфика дискурсивного выражения производственной и общественно-политической тематики обусловлена не столько ее когнитивной неактуальностью, сколько особенностями восприятия соответствующих сторон жизни — сквозь личную сферу носителя народной речевой культуры. Об этом свидетельствует, в частности, эмоциональное согласование фрагментов производственной и общественно-политической тематики с содержанием основных текстовых фрагментов. Негативно оцениваемые факты производственной и общественно-политической жизни обычно включаются в повествования трагического содержания. Различные тематические фрагменты, объединяемые говорящим в один микротекст, очевидно, образуют в его сознании единое эмотивно-смысловое пространство, что объясняет немотивированные, казалось бы, тематические переходы. Например, рассказ о трагической смерти матери включает такие инотематические фрагменты, как упоминание о пятнадцатилетнем периоде тяжелой жизни в доме свекрови и о низкой пенсии, начисленной за длительный — “*всю жизнь*” — тяжелый труд: (5) *...пятнадцать лет / дочка / вот жила и всё терпела // со свекровью жила // вот // вот так вот // и “мама / мама” её / “мама / мама” / всё время / она меня валяла / и всю жизнь вот дояркой проработала // я работала / триста девяносто семь рублей / пенсию / три рубля уж не начислили до четыреста рублей // каталог / ровный // “хватит говорит всё / больше там нечего / говорят / прибавлять” / не начисляют ничего // вот так вот //*. Повествование о трагических событиях в семье брата (смерть брата, гибель племянника, ампутация ноги у жены брата в результате аварии) переходит в краткие рассуждения о бесполезности иметь детей и о невозможности жить и работать в деревне: (6) *...вот живи / живи / для чего живём? вот так разойдусь и говорю / “да Z Z Z я их Z да народила / а?” // разбежись Z да и всё // вёсной летом дочка никого нету / ...а в колхозе работать / за буханку хлеба работал / а щас и буханку хлеба не дают дочка // нету // нету всё // ...вот дают пенсию-то / я вот и подсчитала / нам на двоих только дочка на хлеб // ...и что деревня вот / и сравнили / не нарочно говорят / “сравнили город / с деревней” // одинаково стало дочка // оно и сено купленное / и продукты все купленные //*.

Слияние с личной сферой говорящего проявляется также и при дискурсивном развертывании *религиозной* тематики. Микротемы этой предметной области вербализуются в тесном синтезе с микротемами тематических блоков *частная жизнь* и *дом и домашнее хозяйство*. Например, повествование о том, что старообрядцы (“монашки”) не понимают православных религиозных праздников конкретизируется сообщением о питании рассказчицы во время поста: (7) *...они вот не понимают / что праздники / что / ага/ первый вот / к примеру вот*

я говорю / первый Спас / второй Спас / вот двадцать восьмого Успенье / сейчас вот пост / ну// пост / не умираем -же вот / ну всё есть вон в огороде / где помидорку / огуречек вот я утречком покрошила / луковицу покрошила / картошки покрошила / ну сделала вот салат // дыню сорвала спелую / вот позавтракали / ну? можно терпеть? и яблоки сейчас есть / всё есть / ну / постным маслом помазала // вот двадцать восьмое придет разговенье / пожалуйста / свари / мясо свари там если есть кого там зарежь или чего // а эти / они вообще не понимают / что Троица придет / что Рождество придет / им власть безвластна //.

Ответ на вопрос диалектолога о различии трех Спасов сводится в основном к информации об урожае в текущем году :

(8) Д: *А вот эти Спасы чем различаются? Три Спаса?*

И.: Вот / всё чтобы было / деревьям / всё / было / чтобы об здравии как... они чтобы на будущий год ещё эт родили / давали нам плода // вот видишь / дочка / нынешний год / вот / помидор мало на кустах / огурцов мало // “дам дождя / не дам плода” / видишь // сухое лето / жарко / хорошее было всё / а вот прошёл такой дождик / и всё вон гниёт // всё // вот на сырых местах картошку-то подёргали ботву / а там картошка-то вот какая (#показывает#) и по кулаку и крупнее / / а она вот фишк-фишк и всё // вот // она от влаги / от сырости // она ботва-то вот какая выросла / пронежилась / а Петров-то день / до Петрова-то дня прошла / и Петров / разговляться новой картошкой / а она у нас / поздно садим / только завязалась / а выросла / пошли гляди-ка дожди // то град побил / то половодь была / илом-то всё занесло / ...

Противопоставление православных и иноверцев реализуется через посредство эпизодов бытового характера:

(9) *...они ходят кресты не носят / они не понимают / что такое у них эт есть власть / что за бог / вот // вот вчера эт тут девчонка-то у них / дочь / Оля / Оля // о! (#изображает; смех#) / “старые бессовестные!” меня / кричала орала //;*

(10) *...они не понимают вот что власть // ... эй! глядишь баню затопили // “да нельзя нынче / баню топить / праздник-то какой” // “да-а / это мыться не грех” говорит и всё / понял?;*

(11) *...вот мы православные / мы всё дозволяем // мы дадим и пить / и покормим / и чего кто попросит всё даем // а у кулугур / у них Z особо / чашку они воды не дадут Z //.*

Основные и периферийные тематические блоки различаются степенью конкретизации хронотопа, точнее его темпоральной составляющей. Общим для всех тематических сегментов является отсутствие указания на место описываемых событий. Это объясняется их обычной локализацией исключительно на территории постоянного проживания информанта. Однако в случаях отступления от данной зако-

номерности топос отмечается, например: *...и вызвали меня / я жила в Полисове /*. Что касается показателей темпоральности, то их использование различно в разных тематических блоках. В рамках основных тематических блоков время описываемых событий максимально конкретизировано, в рамках периферийных – определяется весьма приблизительно, посредством таких, например, операторов, как *всю жизнь, раньше – сейчас*. Ср.: (12) *...раньше давали было в сельсовет / от сельсовета давали доски / повезут на пиломару и сделают / а сейчас / всё своё / как хочешь так и поминай // где их тысячи брать? вот так вот //*. См. также примеры (5) и (6).

Значимость темпоральной конкретизации для выявления дискурсивной специфики различных тематических сегментов текста определяется тем, что их использование спонтанно, не инициировано вопросами собеседников. Введение темпоральных конкретизаторов может изменить структуру высказывания. Например, высказывание обобщающего характера (пример 7), содержащее вводное выражение *к примеру вот я говорю*, после использования темпоральных конкретизаторов (*вот двадцать восьмого Успенье / сейчас вот пост*) преобразуется в структуру, описывающую конкретную ситуацию. Кроме того, для основных тематических блоков характерна избыточность хроно-топических показателей, выражающаяся повторением детерминантов со значением места и времени. Ср. примеры из тематического блока *частная жизнь*: (13) *... у меня брат помер / в Москве / в Москве брат помер / и я ездила на эти / на похороны / на похороны // в Москву ездила // это было дело в семидесятом году //*; (14) *... теперь и сын-то / погиб // двадцать семь лет сравнялось // двадцать седьмого июля / вот // двадцать седьмого июля / двадцать семь лет / и двадцать седьмого числа в этом месяце и в этот день / погиб //*.

Создание тематически размеченного машиннообрабатываемого текстового диалектного корпуса позволяет получить принципиально новые сведения о диалекте, открывает перед исследователями возможности изучения диалекта не только как специфической языковой системы, но и как особого культурно-коммуникативного образования. В частности, широкая и узкая тематическая рубрикация диалектных текстов, анализ дискурсивных различий, наблюдающихся в пределах различных тематических блоков текста, записанного от одного информанта, дают возможность определить место и роль той или иной предметной области в когнитивно-дискурсивном поле индивида.

Библиографический список

1. Гольдин В.Е. К проекту текстового диалектологического подфонда Машинного фонда русского языка // Доклады III Всесоюзной конференции по созданию машинного фонда русского языка. М., 1990. С. 92-103.

2. Демешкина Т.А. Теория диалектного высказывания. Аспекты семантики. Томск, 2000. 188 с.
3. Иванцова Е.В. Феномен диалектной языковой личности. Томск, 2002. 312 с.
4. Китайгородская М.В., Розанова Н.Н. Современная городская коммуникация: тенденции развития : (На материале языка Москвы) // Русский язык конца XX столетия (1985-1995). М., 1996. С. 345–381.
5. Китайгородская М.В., Розанова Н.Н. Речь москвичей: коммуникативно-культурологический аспект. М., 1999. 396 с.
6. Летучий А.Б. Диалектные тексты в Национальном корпусе русского языка: принципы представления и разметки // Язык и общество в синхронии и диахронии. Саратов, 2005. С. 44–49.
7. Летучий А.Б. Корпус диалектных текстов: задачи и проблемы // Национальный корпус русского языка: 2003-2005. Результаты и перспективы. М., 2005. С. 215-233.
8. Матвеева Т.В. Непринужденный диалог как текст // Человек – текст – культура. Екатеринбург, 1994. С. 125-140.
9. Русские народные говоры. Звучащая хрестоматия. Южнорусское наречие. М., 1999. 207 с.
10. Савчук С.О. Метатекстовая разметка в национальном корпусе русского языка: базовые принципы и основные функции // Национальный корпус русского языка: 2003-2005. Результаты и перспективы. М., 2005. С. 62-88.
11. Сиротинина О.Б. Тексты, текстoids, дискурсы в зоне разговорной речи // Человек – текст – культура. Екатеринбург, 1994. С. 105-124.

Г.Н. Манаенко

ЗНАЧЕНИЕ И ПЛАН СОДЕРЖАНИЯ ЯЗЫКОВОГО ВЫРАЖЕНИЯ: ЛИНГВОКОГНИТИВНОЕ ИЗМЕРЕНИЕ

Ставропольский государственный педагогический институт

В гуманитарных науках существует разное понимание содержания термина “значение”: если психолог относит к объектам сознания прежде всего “предметные” значения, то для лингвистов – это прежде всего словесные значения, причем, данные значения оказываются соотносительными. В этой связи А.А. Леонтьев пишет: “Каким образом значение слова “стол” связано со значением реального стола, представленному нашему сознанию в виде образа того или иного рода или приписываемого нами такого образа. Как философ я могу ответить: этот образ ни в коей мере не знак, но какой-то вид идеального объекта, спроецированный на реальный объект, слитый с ним субъективно в единое целое (“удвоение” предмета, по Выготскому). Как