

К. А. Маслинский

СОВЕТСКИЙ УЧИТЕЛЬ НА ФОНЕ ШКОЛЬНОЙ ПОВЕСТИ: КОРПУСНАЯ ПЕРСПЕКТИВА

Задача данной статьи — анализ дискурсивного фона, на котором существуют образы учителей в советской школьной повести послевоенного периода. Материалом послужил корпус текстов 37 авторов о школе и школьниках, написанных в 1940–1980 гг. общим объемом 1,8 млн слов. С помощью вероятностного тематического моделирования (LDA) был проведен автоматический анализ содержания эпизодов, в которых упоминаются ключевые слова *директор*, *завуч*, *учитель*, *учительница* и выделены темы, встречающиеся в этих эпизодах значимо чаще или значимо реже, чем в целом по корпусу. По характеру тематического профиля термин *учительница* ярко противопоставлен всем остальным. Среди значимых контрастов, различающих тематические профили терминов: воспитательный дискурс и учебная повседневность, комсомол и пионерия, эмоции и гендер.

Ключевые слова: советская детская литература, школьная повесть, учитель, тематическое моделирование, LDA.

В советской детской литературе послевоенного периода изображение школы и школьной повседневности детей занимает такую долю в тематическом спектре произведений, которую трудно не заметить. В прозаических текстах распространение школьной темы в этот период связано прежде всего с жанровой разновидностью, которую в советской и постсоветской традиции принято обозначать термином «школьная повесть». Этот термин получил распространение в конце 1930-х — начале 1940-х годов, когда в Союзе писателей и в литературной критике стали обсуждаться недостатки в изображении советской школы в детской литературе [Добренко 2013]. Тогда же критики стали формулировать задачу создания «хорошей книги о советской школе» [Литвинов 1935]. Во второй половине 1940-х — начале 1950-х годов задачи по созданию произведений на школьную тему неоднократно обсуждаются на различных высоких коллегиях, включающих не только членов Союза писателей и руководство Детгиза, но и руководство министерства просвещения, ЦК ВЛКСМ, Академии педагогических наук РСФСР [Фатеев 2007]. Участие в дискуссиях высокопоставленных

чиновников, ответственных за образование и воспитание подрастающих советских поколений, недвусмысленно свидетельствует о том, что школьная повесть в этот период была частью советского образовательного проекта. В этом проекте детская литература отвечала за разработку моделей репрезентации школы как социального института и связанных с ней воспитательных процессов.

Школьная повесть как жанровая форма советской детской литературы пережила и пределы сталинской эпохи, и более поздние метаморфозы советской образовательной и культурной политики. К концу советского периода едва ли остаются какие-либо основания говорить о государственном заказе на изображение школы в детской литературе. Однако жанровое определение «школьная повесть», занявшее в позднесталинскую эпоху прочное место в жанровой системе детской литературы, и созданные в этот период образцы текстов надолго обозначили отчетливую тематическую нишу для авторов. Кроме того, государственная установка на изображение советской школы в детской литературе открыла советским литературным критикам и постсоветским исследователям путь для обобщения и оценки этих текстов в школьных категориях: учителей, учеников, воспитательного процесса, сознательной дисциплины и др. Так возникла задача реконструкции образа советской школы по текстам школьных повестей, обратная к изначальной задаче авторов по ее изображению.

При этом специфические черты облика советской школы, сложившиеся в жанре школьной повести, нельзя полностью списать на государственный заказ. Например, повесть Фриды Вигдоровой «Мой класс» может быть прочтена и как вполне конформистская иллюстрация модели авторитарного советского учителя [Литовская 2010], и как реализация новаторской для ортодоксальной советской школы того времени педагогической концепции индивидуального подхода к ученику [Майофис, в печати].

Возможности управлять деталями репрезентации школы в литературных текстах ограничены «разрешающей способностью» инструментов государственного регулирования литературы. Разного рода декларации и призывы к писателям, претендующие на роль творческого задания, формулировались в очень широких тематических категориях. Более тонкая настройка приемлемых и неприемлемых элементов содержания выполнялась на уровне обратной связи: редактурой, оценкой в критике, наградами и цензурой уже созданных текстов. Так, Сталинская премия третьей степени, полученная

Николаем Носовым в 1952 г. за повесть «Витя Малеев в школе и дома», указывала образец, пригодный для тиражирования в жанре, хотя и не выделяла конкретных свойств этого образца, достойных подражания. Отбраковывать и одобрять элементы содержания на уровне фабулы и персонажей позволяла редаKTура и критика. И все же в текстах оставалось достаточно степеней свободы, зависящих от автора и других обстоятельств и не контролируемых непосредственно литературной бюрократией. В частности, менее подконтрольными оказывались все те характеристики образа школы, которые не сфокусированы в одной точке текста, но рассеяны во множестве мелких деталей на уровне языка и изображаемых ситуаций повседневности.

Синтетический образ советской школьной повседневности, складывающийся из всех уровней произведения, можно структурировать с помощью дихотомии фигуры и фона. *Фигурой* в данном случае является «сообщение» автора, соотносимое с задачами произведения, литературной и политической конъюнктурой и легко объяснимое в их терминах. Фигура связана прежде всего с уровнем фабулы, но возможна она только в силу существования *фона*, под которым следует понимать ограничения дискурсивного уровня: как и из какого содержательного материала может строиться высказывание о школе. Мне представляется, что изучение фонового образа школы в советской школьной повести, причем в его наиболее стабильных чертах, — необходимое и важное дополнение, без которого невозможно судить о том, в какой мере школьная повесть служила медиатором догматических установок советского образовательного и воспитательного проекта для читательской аудитории.

Традиционные литературоведческие методы позволяют хорошо работать с фигурой, но выглядят гораздо более спорными при попытке описания фона. Подбор знаковых эпизодов, типовых деталей, построение собирательных образов требуют охвата значительно большего количества произведений и неизбежно подразумевают очень высокую степень избирательности при построении базы для обобщений [см. например: Сергиенко 2005]. Альтернативой, позволяющей отчасти разрешить названные проблемы (несомненно, создав новые), служит корпусный метод, предполагающий сплошной количественный анализ текстов произведений. Подобные подходы к литературному процессу как корпусу текстов в последние годы получают все большее распространение и иногда объединяются под общим именем «макроанализ» [Jockers 2013].

В данной статье поставлена относительно узкая задача — анализ дискурсивного «фона» в изображении советских педагогов в школьной повести. Выбор объекта объясняется тем, что идеологически верное изображение учителей было одной из задач школьной повести, эксплицитно сформулированной в позднесталинский период [см. Майофис, в печати]. Не случайно, что анализ образов учителей на уровне «фигуры» присутствует во всех процитированных выше работах о советской школьной повести, а также и в работах западных советологов, видевших в школьной повести нормативную репрезентацию образа учителя, одобренную государством [Muckle 1979].

Ключевая идея предлагаемого мной метода анализа «фона» — оттолкнуться от самых общих лексических категорий, обозначающих школьных педагогов: *директор, завуч, учитель, учительница*. Употребление любого из этих слов в тексте школьной повести можно рассматривать как индикатор того, что в данном эпизоде актуализирована социальная позиция педагога в рамках школьной системы или категория «учитель» в целом. Важно подчеркнуть, что речь идет об отборе эпизодов для анализа именно по лексическому принципу, а не по присутствию в них персонажей-учителей. Анализ содержания контекстов, в которые вписана категория «учитель» в жанре школьной повести, в данном подходе служит инструментом для реконструкции общей для всех рассмотренных текстов семантики этой категории и прагматики ее употребления в школьной повести.

Материалом для данной работы послужила относительно небольшая электронная коллекция текстов школьных повестей 1940-х — 1980-х годов. Нужно оговориться, что при отборе текстов я не придерживался никакого узкого определения жанра школьной повести, а трактовал эту категорию прежде всего тематически, как тексты, в которых значительное внимание уделяется описанию школы и школьной повседневности. В результате в коллекцию попали тексты для младших и старших подростков, описывающие начальную, среднюю и старшую школу, а также некоторые разновидности «каникулярной повести». Разнородный состав коллекции позволяет анализировать концептуализацию советской школы на самом общем уровне. Всего в рассмотренный корпус текстов вошли произведения 37 авторов, общим объемом порядка 1,8 млн слов¹.

Корпусный поход подразумевает не только систематическое рассмотрение большого количества текстов, но и обусловленное ростом

объема анализируемого материала и общей идеологией макроанализа смещение акцента с медленного чтения к автоматизированному количественному анализу. Соответственно, поставленная в данной работе проблема анализа семантики категории «учитель» в рамках корпусного подхода требует формализации и решения следующих задач: сегментация текста на эпизоды, автоматизированный анализ содержания эпизода, выявление элементов содержания, характерных и нехарактерных для эпизодов, где присутствует категория «учитель».

Единица анализа контекста в данной работе — эпизод. Поскольку автоматическое определение границ эпизодов в литературном тексте — задача, на сегодняшний день не имеющая универсального решения, был выбран упрощенный формализованный подход к сегментации текста на эпизоды. Все тексты были разбиты на фрагменты объемом около 500 слов², в сумме в коллекции получилось 5 946 фрагментов (эпизодов). Размер фрагмента подобран эмпирически с расчетом, чтобы в среднем в его рамки с запасом укладывалась непрерывная сцена — диалог, описание и т. п. Границы выделенных таким образом эпизодов не всегда совпадают с границами глав и другими разделами текста и часто захватывают смежные сцены. Однако преимущество выбранного подхода в том, что он позволяет учитывать содержание не только непосредственной сцены, но и ближайшего смежного контекста. Отношения смежности такого рода не находятся в фокусе внимания читателя и даже, возможно, автора, но при накоплении достаточной статистики они могут вскрыть неочевидные, на первый взгляд, синтагматические ограничения жанра.

В качестве основного инструмента для автоматизированного анализа содержания эпизодов использован метод вероятностного тематического моделирования, основанный на так называемом латентном размещении Дирихле (LDA, Latent Dirichlet Allocation) [Blei, Ng, Jordan 2003]. Смысл этого метода в том, что вероятностный алгоритм, используя только информацию о совместной встречаемости слов в текстах (в данной работе — в эпизодах), распределяет всю лексику коллекции на заранее заданное исследователем количество групп — «тем» (topics). «Темы» взяты в кавычки, так как это понятие является лишь довольно грубым статистическим приближением тематической классификации, выполняемой человеком. Несколько упрощая, можно представить «тему» как упорядоченный список слов, отсортированный по убыванию вероятности встретить слово

в рамках темы. Начало этого списка содержит самые характерные слова темы. Хотя алгоритм не располагает никакой информацией о семантике слов и даже о языке текстов, эмпирические результаты его применения к разным коллекциям текстов показывают, что многие из полученных «тем» группируют лексику из одного семантического домена или лексику, характерную для определенного типа дискурса. При том что алгоритм не предлагает никакого общего обозначения для таких тем, исследователь имеет возможность интерпретировать полученные списки слов и назначить каждой теме общую метку, характеризующую ее содержание. Например, просматривая следующий список наиболее характерных слов одной из тем, выделенных алгоритмом в корпусе школьных повестей: *глаз, лицо, губа, рука, взгляд, бровь, волос, голос, улыбка, нос, лоб, взглядывать, щека, подымать, темный, плечо, строгий, широкий, рот* — нетрудно определить, что эта группа лексики сформировалась за счет совместной встречаемости в портретных описаниях, часто связанных с упоминанием взгляда. В результате такой интерпретации можно присвоить теме метку *портрет*. Метод LDA получил в последнее десятилетие широкое распространение в социальных и гуманитарных исследованиях, в том числе и применительно к анализу корпусов литературных текстов [см. например: Blei, 2012; Jockers, Mimno 2013; Rhody 2012; Tangherlini, Leonard 2013, и мн. др.].

Главный параметр алгоритма LDA, задаваемый пользователем — количество тем, которые нужно сформировать. Обычно при меньшем количестве темы получаются более обобщенные, при большем — более частные. Существуют разные методики выбора оптимального количества тем, но одним из ведущих критериев остается прозрачность интерпретации получившихся тем [Chang et al. 2009]. В данной работе было проведено разбиение коллекции на 100, 125, 150, 175, 200, 225, 250 тем. Наилучшей моделью с точки зрения интерпретации оказалось разбиение на 225 тем, которое и было выбрано для анализа. Далее всюду рассматривается только это разбиение. Для получения более качественного результата тематического моделирования была проведена предварительная подготовка текстов: все слова текста были лемматизированы (приведены к начальным формам), из текстов были удалены стоп-слова, не несущие информации о тематике текста (наиболее частотные предлоги, союзы, местоимения), а также имена всех действующих лиц. Последний прием необходим для того, чтобы избежать

искусственной группировки лексики на основании совпадения имен персонажей [Jockers, Mimno 2013]. Всем темам автором были присвоены метки из одного-двух слов, характеризующие их содержание.

В полученной тематической модели школьных повестей наибольший интерес в рамках поставленной задачи представляют «сквозные» темы, обнаруживающиеся в значительном количестве текстов. Такие темы можно объединить в несколько содержательных групп: описание человека (*портрет*³, *карман*, *платье* и др.), персонажи (*мама*, *девочка*, *семья*, *бабушка* и др.), эмоции (*смех*, *счастье*, *плач*, *обида*, *страх* и др.), сюжетные ситуации (*драка*, *сон*, *гости*, *болезнь* и др.), локусы (*дверь*, *дом*, *лес*, *двор* и др.), природа (*небо*, *зима*, *погода* и др.), предметный мир (*мебель*, *письмо*, *машина*, *телефон* и др.). Немалую часть составляют «школьные» темы: *воспитание*, *кабинет директора*, *класс*, *учиться*, *тетрадь*, *экзамен*, *сочинение* и др. Особняком стоят темы, объединившие в основном не предметную, а нарративную лексику, и, по всей видимости, характеризующие разновидности повествовательного дискурса. Эти темы получили условные названия по автоматически отнесенным к ним фразам: *пройти мимо*, *махнуть рукой*, *покачать головой* и др.

Лексемы, обозначающие учителей, в корпусе довольно частотны. Всего насчитывается 627 фрагментов, где хотя бы один раз встречается слово *директор* (в значении ‘директор школы’), *завуч* — 134 фрагмента, *учитель* — 793 и *учительница* — 797. В части эпизодов присутствуют одновременно два или более из этих терминов, например, *директор* и *завуч*, *завуч* и *учитель* и т. п.⁴ В подобных случаях эпизод учитывался в подсчетах тематики несколько раз, как контекст для каждого из упомянутых в нем терминов. Для анализа гендерной специфики термины *учитель* и *учительница* жестко разделены: для *учителя* учитываются только такие эпизоды, где не упоминается слово *учительница*.

Базой для анализа жанрового «фона» каждого из рассматриваемых терминов стал его суммарный контекст — все эпизоды (фрагменты), где хотя бы один раз упомянут данный термин. Алгоритм тематического моделирования не только группирует лексику, но и вычисляет, какие из тем и в какой пропорции представлены в каждом из эпизодов. С помощью этих данных можно оценить, какие темы чаще или реже присутствуют в контекстах данного

термина, чем в целом по корпусу. Выявленные таким образом темы показывают, с какими семантическими зонами преимущественно ассоциируется рассматриваемая категория, а каких зон она избегает. Учитывались только те темы, которые по оценке алгоритма занимали не менее 5% объема эпизода (что при эпизоде объемом в 500 слов равно 25 словам). Этот порог введен для снижения статистического шума. Также из рассмотрения были исключены все темы, специфические для отдельных произведений коллекции⁵ и очень редкие темы, встречающиеся менее, чем в 1% фрагментов корпуса. Для оценки статистической значимости (то есть исключения гипотезы, что повышенная или пониженная частотность появления темы может быть целиком объяснена случайностью) использовался критерий Фишера⁶.

Если выписать все темы, чаще встречающиеся во всей совокупности эпизодов с упоминанием директора, завуча или учителя/учительницы, чем в целом в школьных повестях, результат не будет особенно неожиданным. Практически полностью список состоит из очевидно «школьных» тем: *учительница, кабинет директора, комсомол, учительский стол, класс, воспитание, двойка/отметки, иметь право, тетрадь, экзамен, пионеры, урок, выбор профессии, поступок, приглашать, девочка, институт, ребята/шум*. Темы в приведенном здесь списке упорядочены по убыванию статистической значимости: чем ближе к началу списка, тем сильнее заметна повышенная частотность темы в «учительских» эпизодах по сравнению с ее средней частотностью в коллекции. Отдельного комментария заслуживают, пожалуй, несколько моментов.

Во-первых, темы *комсомол* и *пионеры* оказываются значимо связаны в школьной повести с упоминанием учителей. Хотя доля эпизодов с упоминанием комсомола и пионерии почти одинакова (7% и 6,2% соответственно), связь комсомола с упоминанием учителей значительно теснее, так как эта тема намного реже, чем тема пионерии, встречается там, где учитель не упомянут. Во-вторых, в списке присутствует несколько тем, характеризующих, скорее, дискурсивные феномены, чем предметную область. Сюда относятся тема *иметь право* (наиболее характерные слова темы: родитель, право, отношение, девчонка, иметь, против, завуч, разговор, взгляд, вопрос), *поступок* (считать, поступок, решение, прежде, помогать, характер, поступать, год, нужный, принимать), *приглашать* (ваш, приглашать, товарищ, обращаться, разговор,

время, находиться, несколько, просить, продолжать). Ядро всех трех тем, как мне представляется, составляет лексика, связанная с описанием разного рода воспитательных разговоров, в том числе обсуждения поступков. Впрочем, в этой интерпретации без опоры на тексты так велика доля интуитивной реконструкции, что следует рассматривать основанный на них вывод как предварительную гипотезу.

Более интересные результаты можно получить, если выписать темы, связанные с упоминанием каждой из школьных ролей в отдельности. Статистика, отражающая распределение значимых тем в эпизодах с упоминанием директора, завуча, учителя и учительницы, приведена в таблице 1. Прочерки в таблице обозначают, что доля темы среди эпизодов с данным словом существенно не отличается от ее доли в целом по корпусу. Темы в таблице сгруппированы в блоки по характеру их распределения по школьным ролям. Таблица показывает, что тематические портреты всех четырех рассмотренных терминов различны. Наиболее заметно различие между термином «учительница» и всеми остальными, которое формируется главным образом за счет блока тем, характеризующих учебную повседневность и общение со школьниками (*девочка, пионеры, ребята/шум, тетрадь*). В то же время, у термина «учительница» отсутствует статистически значимая связь с темами «дискурсивного» блока (*иметь право, поступок, приглашать*). Учительнице противопоставлены директор и завуч, очень похожие по тематическим спектрам, включающим общешкольные и воспитательные темы (*кабинет директора, комсомол, воспитание*) и уже упомянутые дискурсивные темы. Различие между ними в том, что директор связан с классом, но не с уроком, а завуч — наоборот. Учитель занимает в тематическом плане промежуточное положение между ними и учительницей: воспитательные темы объединяют его с директором и завучем, образовательные (*учительский стол, двойка/отметки*) — с учительницей. Кроме того, учитель — единственный термин, значимо связанный с тематикой профессиональной ориентации (*выбор профессии*). Количественные различия в степени выраженности темы в общем для всех терминов блоке (*кабинет директора, комсомол, воспитание*) позволяют выстроить градиацию терминов по степени убывания связи с общешкольной и воспитательной тематикой: завуч—директор—учитель—учительница.

Таблица 1

Доля фрагментов (в процентах), в которых присутствует названная тема. В колонках — процент от общего числа фрагментов с указанным ключевым словом. Последняя колонка — общий процент фрагментов в корпусе, в которых присутствует данная тема.

тема	директор	завуч	учитель	учительница	в корпусе
кабинет директора	28,1	44,0	15,1	6,3	4,1
комсомол	9,9	10,5	8,3	5,4	3,2
воспитание	14,5	19,4	17,6	9,0	6,8
иметь право	5,1	26,9	6,6	—	2,4
поступок	8,5	14,2	9,3	—	5,3
приглашать	—	6,0	3,3	—	1,9
класс	9,1	—	8,4	12,4	5,3
урок	—	9,0	6,5	5,3	3,3
учительница/урок	4,3	8,2	—	12,7	1,9
учительский стол	—	—	13,3	22,9	7,4
двойка/отметки	—	—	6,5	10,3	3,9
девочка	—	—	—	13,2	4,9
пионеры	—	—	—	9,7	4,0
ребята/шум	—	—	—	6,6	4,3
тетрадь	—	—	—	3,5	1,3
выбор профессии	—	—	5,7	—	2,4
институт	—	—	—	—	3,4

Тематические портреты каждого из терминов можно дополнить значимыми темами, отсутствующими в общем списке. Среди них обращают на себя внимание темы, содержащие лексику описания эмоциональных состояний. Так, тематический портрет директора может быть дополнен темой *страх* (наиболее частотные слова темы: бояться, страшный, голос, правда, палец, трус, слышать, разговор, кричать, смелый), а тематический портрет завуча и учителя — темой *неприятные переживания* (слово, плохой, душа, мысль, чувство, чувствовать, время, случаться, стараться, часто). Дополнительными штрихами к портрету термина «учитель» становится тема *наука*, а к термину «учительница» — темы *математическая задача* и *гости* (здравствуйте, гость, рука, веселый, сторона, стол, место, плечо, показывать, рад).

Менее очевидной для читателя школьной повести, но не менее информативной для описания синтагматики жанра стороной тематического портрета термина являются темы, которые появляются в его контексте *существенно реже*, чем в целом по корпусу. За счет таких тем можно оценить контексты, которые в рамках жанра максимально далеки от семантики интересующих нас школьных ролей или даже находятся с ними в отношении дополнительной дистрибуции. Список значимо более редких тем во всей совокупности эпизодов с упоминанием директора, завуча и учителя/учительницы включает 34 темы, охватывающие многие внешкольные топосы советской детской литературы. Сюда относятся семья (темы *отец/мать, папа/мама*), еда и покупки (*еда, застолье/выпивать, сладости, деньги/купить*), детские занятия (*каток, игры, драка*), загородные локусы (*река, лес, море, деревня, огород*), транспорт (*поезд, самолет*) и пейзаж (*небо, освещение*). Даже этот неполный список показывает, что упоминание категорий директор/завуч/учитель в школьной повести довольно жестко разведено с описанием практически любых внешкольных контекстов, то есть, по существу, эти термины не выходят за рамки описания школы. Этот результат может показаться не слишком неожиданным, однако он накладывает некоторые ограничения на возможные спекуляции о всепроникающем присутствии дискурса школы и школьной дисциплины в советской школьной повести [ср. Добренко 2013].

Для каждого из терминов есть и свой специфический набор тем, существенно более редких в его контексте. Если в случае директора список редких тем в целом воспроизводит общий набор для всей совокупности упоминаний школьных педагогов, то к портрету завуча, учителя и учительницы с помощью специфических недостающих тем можно добавить любопытные детали. Самый яркий результат — полное отсутствие в эпизодах, содержащих слово «завуч», темы *смех*, которая в целом по корпусу не так редка и встречается в 5,3% всех эпизодов. Этот результат тем замечательнее, если принять во внимание, что автоматически выделенные фрагменты часто захватывают смежные эпизоды, то есть в пределах 500 слов в контексте упоминания завуча в советской детской литературе никто не смеется.

Для термина «учитель» среди специфических редких тем обнаруживается *девочка* (2,8% против 5,0% в целом по корпусу), что демонстрирует неожиданную гендерную избирательность в контекстах этого, казалось бы, не маркированного в гендерном отношении слова. В контексте упоминания учителя и учительницы

не характерна тема драки, или, возможно, правильнее будет сформулировать это так, что при описании драки не характерно упоминание слов «учитель» и «учительница». Наконец, среди специфически редких для «учительницы» тем обнаруживается тема *автомобиль* (машина, шофер, дорога, ехать, грузовик, колесо, кабина, рядом, кузов, поехать).

Возвращаясь к поставленному в начале статьи вопросу о дискурсивном «фоне», на котором вырисовываются образы советских учителей в школьной повести, следует просуммировать общие тенденции в контекстах соответствующих терминов. По большому счету, искомым фоном оказалась советская школа в разных более и менее очевидных ее повседневных проявлениях, а присутствующие в школьных повестях внешкольные контексты не внесли в этот фон практически никакого вклада. Говоря о дискурсивном фоне к отдельным терминам, обозначающим советских педагогов (директор, завуч, учитель и учительница), уместно вспомнить, что обозначаемые ими роли строго упорядочены по внутришкольному статусу⁷. Иерархия школьных статусов задает внешнюю ось, относительно которой можно выстраивать и тематические различия в контекстах этих статусов в школьных повестях. Вдоль этой иерархической оси обнаружилось несколько интересных тематических контрастов.

Прежде всего, это контраст школьного воспитания и учебы. Воспитательные темы отчетливо тяготеют к верхней части иерархической шкалы (директору и завучу), а лексика учебной повседневности жестко закреплена за ее нижней частью (учителем и в особенности учительницей). Обнаруженная дистрибуция расходится с настойчиво педалировавшейся позицией советской педагогики о единстве учебного и воспитательного процессов. По меньшей мере, в школьной повести вместе с понижением статуса педагога падает и градус воспитательного дискурса. Трудно предположить, что в советской школьной повести учительница не воспитывает, однако, по данным корпусного тематического анализа, ее воспитательная лексика должна быть во многом иной, чем у директора и завуча.

Контраст комсомола и пионерии тоже отчетливо выстраивается вдоль иерархической оси школьных статусов. Более выраженная связь комсомола с вершиной иерархической шкалы легко объяснима его идеологической и воспитательной ролью в школе и ведущей ролью по отношению к пионерской организации. Любопытно в этом контрасте, скорее, отсутствие связи между темой пионерии и всеми статусами кроме самого низкого — учительницы. Это обнаруживает

имплицитную дискурсивную границу для изображения воспитательной роли пионерии в детской литературе.

Раздельный анализ контекстов для терминов «учитель» и «учительница» позволил вскрыть, как гендерное неравенство в среде советских учителей транслировалось на уровне репрезентации их символического капитала в школьной повести. Только учитель, но не учительница оказывается связан с темами, обосновывающими профессиональный статус, — связь с наукой и роль при обсуждении выбора профессии. В то же время учительница оказывается в гораздо большей степени связана с учебной повседневностью и, шире, вообще остается единственным из рассмотренных терминов, хоть сколько-то значимо связанным с темами из мира детей. В контекстах, окружающих упоминание учительницы, встречаются и шумящие ребята, и тема пионерии, и хождение в гости. Таким образом, только самый низкостатусный советский педагог смыкается в синтагматике жанра школьной повести с внешкольным миром, в то время как более высокостатусные школьные педагоги остаются либо выше, либо в стороне от топосов детской повседневности.

Способ работы вероятностного алгоритма тематического анализа, не делающего априорных различий между словами, позволил обнаружить контрасты не только в предметном, но и в дискурсивном контексте разных школьных статусов. Так, дискурсивные темы позволяют построить эмоциональную ось, на которой директор, завуч и учитель противопоставлены учительнице. Верх этой шкалы связан с разного рода отрицательными эмоциями: директор — со страхом, завуч и учитель — с неприятными переживаниями. Кроме того, в контекстах, где упоминается завуч, отсутствует смех. Внизу шкалы, в контекстах с упоминанием учительницы, значимо не выделяются ни положительно, ни отрицательно маркированные эмоциональные темы. На аналогичные сегменты делится иерархическая шкала при рассмотрении дискурсивных воспитательных тем (*иметь право, поступок*), которые значимо связаны только с директором, завучем и учителем. Причем наиболее выражена эта связь не с директором, а именно с завучем. Этот факт и эмоциональная характеристика этой роли подталкивают к заключению, что именно понятие «завуч» оказывается в школьной повести наивысшим воплощением дискурса школьной дисциплины и точкой приложения эмоционально напряженного взаимодействия со школой. Примечательно, что в этой позиции оказывается не директор школы, который, по всей видимости, в рамках школьной повести в максимальной степени олицетворяет

советский образовательный проект, и поэтому авторы избегают упоминаний директора в слишком рискованных контекстах.

В заключение в лучших традициях советской самокритики отмечу, что не следует переоценивать надежность обобщений, сделанных в этой статье. Прежде всего, они построены на слишком небольшом, довольно случайно подобранном корпусе текстов и при условии очень широкой жанровой трактовки школьной повести. Такой выборки текстов может быть достаточно, чтобы кинуть беглый взгляд на самые общие черты в содержании послевоенной советской детской литературы. Но в действительности эта литература имеет сложную внутреннюю структуру — и хронологическую, и стилевую, и, вероятно, идеологическую, даже если мы ограничим рассмотрение более или менее узко понятым жанром школьной повести. Чтобы увидеть в корпусном масштабе эти различия и временные тенденции, нужен гораздо больший и, что важнее, систематически составленный корпус текстов. Тем не менее, в защиту полученных здесь результатов могу добавить, что главное достоинство автоматизированных количественных методов анализа текстов не столько в неожиданных новых результатах, сколько в новых вопросах и направлениях поиска, которые удастся обозначить с их помощью.

Примечания

¹ А. Алексин, М. Бременер, И. Василенко, Е. Велтистов, Е. Верейская, В. Голявкин, Л. Давыдычев, И. Дик, В. Добряков, Н. Дубов, Ю. Ермолаев, В. Железников, Е. Ильина, Л. Исарова, И. Карнаухова, В. Киселев, М. Коршунов, А. Кузнецова, А. Лиханов, Г. Матвеев, Л. Матвеева, Г. Михасенко, А. Мусатов, Л. Никольский, Н. Носов, В. Осеева, И. Пивоварова, Т. Поликарпова, Г. Полонский, М. Прилежаева, Ю. Сальников, Л. Симонова, Г. Скребицкий, Н. Соломко, В. Фролов, Е. Шварц, Г. Щербакова.

² Граница фрагмента проводилась после конца того абзаца, в котором объем достиг 500 слов, чтобы избежать обрывов текста на полуслове.

³ Здесь и далее при указании на тему будут использоваться метки, присвоенные им в результате интерпретации.

⁴ Количество пересечений разных терминов в рамках одного эпизода превышает порог, объяснимый случайным совпадением, для всех пар терминов, кроме заучу — учительница. В качестве инструмента оценки порога случайности использован точный тест Фишера [Baroni, Evert 2009].

⁵ Тема считалась специфической для произведения, если все фрагменты, где эта тема занимает не менее 5% объема, относились к одному и тому же произведению. Всего таких тем оказалось 41 из 225.

⁶ Точный тест Фишера выбран, так как он подходит для небольших выборок и широко применяется в аналогичных случаях в корпусной лингвистике [Baroni, Evert 2009]. Рассматривались только результаты на уровне статистической значимости 0,01. Относительно высокий порог значимости установлен для того, чтобы ограничить

рассмотрение только наиболее существенными для соответствующих фрагментов тематическими контекстами. В силу высокой степени неточности как в определении границ эпизодов, так и автоматическом выделении тем, при рассмотрении более слабых эффектов велик риск получения значимых результатов из-за разного рода случайных погрешностей.

⁷ Термин «учительница» может быть отнесен на более низкую ступеньку, чем «учитель», так как акцентирует менее престижный гендер и часто связан с менее статусной начальной школой.

Исследования

Добренко Е. А. «...Весь реальный детский мир» (школьная повесть и «наше счастливое детство») // «Убить Чарскую...»: Парадоксы советской литературы для детей 1920-е — 1930-е гг. / под ред. М. Р. Балиной, В. Ю. Вьюгина. СПб.: Алетей, 2013. С. 189–230.

Литвинов В. Дореволюционная и советская школа в детской литературе // Детская литература. 1935. №. 12. С. 1–8.

Литовская М. А. Школьная повесть как инструмент анализа повседневности советской школы // Антропология советской школы: Культурные универсалии и провинциальные практики. Пермь: Пермский гос. ун-т, 2010. С. 278–291.

Майофис М. Л. Предвестия «оттепели» в советской школьной политике позднего сталинского времени // Острова утопии: Педагогическое и социальное проектирование послевоенной школы (1940–1980-е). М.: Новое литературное обозрение, в печати.

Сергиенко И. А. Школьный мир на страницах школьной повести 1920–1980-х годов // Литературные явления и культурные контексты: материалы коллоквиума молодых ученых-гуманитариев Санкт-Петербурга и Даугавпилса / под ред. М. Л. Лурье. СПб., 2005. С. 59–67.

Фатеев А. В. Сталинизм и детская литература в политике номенклатуры СССР. М.: МАКС Пресс, 2007.

Baroni M., Evert S. Statistical methods for corpus exploitation // Corpus linguistics: An International Handbook. Vol. 2. Berlin, New York: Walter de Gruyter, 2009. P. 777–802.

Blei D. Topic modeling and digital humanities // Journal of Digital Humanities. 2012. Vol. 2, no. 1. P. 8–11.

Blei D. M., Ng A. Y., Jordan M. I. Latent dirichlet allocation // The Journal of Machine Learning Research. 2003. Vol. 3. P. 993–1022.

Chang J. [et al.] Reading Tea Leaves: How Humans Interpret Topic Models // Advances in Neural Information Processing Systems. 2009. Vol. 22. P. 288–296.

Jockers M. L. Macroanalysis: Digital methods and literary history. University of Illinois Press, 2013.

Jockers M. L., Mimno D. Significant themes in 19th-century literature // Poetics. 2013. Vol. 41, no. 6. P. 750–769.

Muckle J. Images of the Teacher in Stories for Soviet Children // Compare: A Journal of Comparative and International Education. 1979. Vol. 9, no. 1. P. 59–64.

Rhody L. Topic modeling and figurative language // Journal of Digital Humanities. 2012. Vol. 2, no. 1. P. 19–35.

Tangherlini T. R., Leonard P. Trawling in the Sea of the Great Unread: Sub-corpus topic modeling and Humanities research // Poetics. 2013. Vol. 41, no. 6. P. 725–749.