

цистики (30 852 670 из 44 014 121 словоупотреблений) – здесь присутствуют в первой двадцатке: *я* – 16 позиция 44 014 121 единиц, *мы* – 17 позиция и 328 822 единиц, *ён* – 20 позиция и 292 791 единиц (см. Беларускі N-корпус).

Доступные для анализа данные относительно частеречного состава ресурса *Беларускі N-корпус* требуют дальнейшей обработки по причине «неснятой омонимии» в корпусе: например, определить пропорциональное соотношение союза *што* и местоимения *што* (рус. *что*) сегодня можно только приблизительно: контексты корпусного конкорданса свидетельствуют об относительно небольшом количественном преобладании союза *што* над местоимением *што*:

Ён пабачыў, што я хачу падняцца, а не падымуся;

Мы насталі: думалі спачатку – немцы, зірнулі – пазналі, што гэта ідуць партызаны;

Тое, што адбывалася з Вялікімі Прусамі, дзве гэтыя жанчыны бачылі, перажылі кожная па-свойму;

Пытанне: – А пра што яны ўвечары гаварылі з людзьмі? (Алесь Адамовіч, Янка Брыль, Уладзімір Калеснік. Я з вогненнай вёскі, 1975).

Таким образом, приведенные примеры корпусный менеджер ресурса *Беларускі N-корпус* сопровождал, к сожалению, недостаточно дифференцированным в отношении частеречной принадлежности единиц метаописанием. Аналогично, зачастую, организовано метаязыковое сопровождение текстов в большинстве лингвистических корпусов, при этом, например, в широко известном НКРЯ омонимия снята только для сравнительно незначительного объема (в сумме около 6 млн словоупотреблений) текстов – работа выполнена сотрудниками проекта «вручную» (см. Национальный корпус русского языка). Таким образом, качество исследования КОД пока в значительной степени обусловлено задействованностью в пред- и постобработке текстов человека.

Тем не менее, *индекс семантической*, представляя базовую информацию о конкретном параметре семантической идентичности текста, позволяет проводить достаточно широкие обобщения. Практика свидетельствует, что программными средствами можно обеспечить выполнение целого ряда алгоритмов лингвистической обработки текстов. Разработка узкоспециальных изначально задач в прикладном ключе часто характеризуется перспективностью, форматностью и возможностями включения в комплексные исследования.

Н. А. Богданова

ФОРМАЛИЗОВАННОЕ ПРЕДСТАВЛЕНИЕ ВНУТРИЖАНРОВОГО ТЕКСТОВОГО СХОДСТВА НА ЛОГИКО-СЕМАНТИЧЕСКОМ УРОВНЕ

Формализация логико-семантической организации текста с последующим сопоставлением полученных моделей опирается на следующие исходные предпосылки.

Во-первых, в теории искусственного интеллекта допускаются некоторые огрубления, одним из которых является абсолютизация момента устойчивости психических процессов, т.е. сознание рассматривается не как процесс, а скорее как результат определенного развития биологической системы. В связи с этим приоритетное значение получает статичный аспект, ибо движение может пониматься как совокупность различных состояний моделируемого объекта.

Во-вторых, текст воспринимается, прежде всего, как носитель определенной информации. Моделирование изучаемого объекта предполагает фокусировку на значимой части объекта при одновременном выключении информационного шума. При таком подходе «за кадром» остается область коммуникативной цепочки, которая связана с личностью коммуникатора и коммуниканта, причинами и целями, условиями продуцирования текста. Они могут привлекаться для трактовки полученных результатов на более позднем этапе.

В-третьих, важнейшим фактором, влияющим на построение текста, признается давление системы. Так, исследуя особенности жанра, мы исходим из того, что «из конкурирующей пары жанр – личность автора более высокое место в иерархическом ряду занимает жанр, отношение между ними всегда разрешается в пользу жанра». Исследователи справедливо указывают на наличие определенной интуитивной жанровой рефлексии (Т. В. Шмелева), на существование определенных стереотипов, когнитивных типов (А. Г. Баранов).

Алгоритм формализованного описания содержания текста включает следующие этапы.

1. Построение для каждого текста таблицы основного статического содержания текста (ТОСС). ТОСС в сжатом виде передает основную информацию текста. В ТОСС входят обозначения действующих лиц и другие элементы номинационных цепочек, обозначающие персонажей, топонимы, определяющие художественное пространство произведения, темпоральные указатели, которые свидетельствуют о развитии, продвижении действия. ТОСС строится автоматически на основе алфавитно-частотного словаря и редактируется с целью устранения информационного шума. В формализованном виде в ТОСС входят главные и второстепенные субъекты (ГС, ВС), главные и второстепенные объекты (ГО, ВО), главные и второстепенные действия (ГД, ВД), главное и второстепенное время (ГВ, ВВ), главное и второстепенное место (ГМ, ВМ), главные и второстепенные признаки (ГП, ВП). Таким образом, процесс составления ТОСС включает обработку алфавитно-частотного словаря специальной компьютерной программой Unify, распределение опорных слов в ТОСС с учетом их референтной отнесенности и роли в сюжетодвижении, дополнение списка опорных слов полными словами из заглавия текста (в случае, если они не были выделены автоматически), удаление из списка ТОСС слов, содержащихся в «отрицательном» словаре, редукция списка ТОСС до 10 единиц на основе критериев частотности. Пример ТОСС одного из исследованных текстов представлен в таблице.

Основное статическое содержание текста В. Борхерта «Das Brot»

Тип опорного слова	Субъекты	Объекты	Действия	Время	Место	Признак
ГС ₁ ГС ₂ ГО ₁ ГД ₁ ГД ₂	SIE ER	BROT	SEHEN SAGEN			
ВО ₁ ВД ₁ ВВ ₁		BROT- TELLER	DENKEN		NACHT	

2. Создание метаязыка формализованного описания текста. Метаязык создается с использованием ТОСС, рематических доминант текстовых фрагментов и принадлежности фрагментов к одному из пяти речевых регистров: изобразительному, информативному, генеритивному, волюнтивному или реактивному (по Г. А. Золотовой).

3. Построение логико-семантических формул текстов (ЛСФТ) с использованием вышеописанного метаязыка. ЛСФТ представляет собой линейную последовательность микротем, организованных в регистровые блоки, которые отражают авторское членение текста на абзацы. Ниже приведен пример ЛСФТ рассказа В. Борхерта «Die Küchenuhr»:

0. Ио-1 (ИоМ001) 1. Ио-Ип-11 (ИпМ005 + ИоМ013) 2. Ио-Ип-19 (ИпМ002 + ИоМ018) 3. Ип-6 (ИпМ006) 4. Ио-Ип-20 (ИпМ005 + ИоМ018) 5. Ио-Ип-20 (ИпМ005 + ИоМ018) 6. Ип-2 (ИпМ002) 7. В-4 (ВМ009) 8. И-Р-1 (РМ001 + ИпМ006) 9. Р-5 (РМ022) 10. И-Инф-Р-11 (РМ001 + ИпМ005 + ИоМ019 + ИнфоМ003) 11. И-Инф-Р-12 (ИпМ006 + РМ004 + ИнфпМ002 + ИнфоМ003) 12. И-Инф-8 (ИпМ002 + ИнфпМ002 + ИнфоМ005) 13. Инфо-Инфп-1 (ИнфпМ002 + ИнфпМ005 + ИнфоМ004) 14. И-Инф-9 (ИоМ005 + ИпМ002 + ИпМ006 + ИнфоМ004) 15. В-И-10 (ИоМ005 + ИпМ012 + ВМ012) 16. И-Инф-Р-13 (ИпМ013 + РМ023 + ИнфоМ005) 17. Ип-4 (ИпМ004) 18. И-Инф-10 (ИпМ006 + ИнфоМ003) 19. Ип-10 (ИпМ009 + ИпМ005)

После формализации содержания исследуемых текстов сравнительному анализу подлежат ЛСФТ, их объем, состав, порядок следования микротем. В частности, для жанра немецкого короткого рассказа сравнительный анализ формул позволил выявить следующие общие черты текстов:

1) сравнительно небольшое число микротем информативного регистра, т.е. таких, которые представляют информацию, непосредственно не участвующую в сюжетодвижении;

2) для сильных позиций текста типичны микротемы волонтивного и реактивного регистров. При обращении к вербальной составляющей микротем было отмечено наличие обратнаправленных (в начальных позициях) и прямонаправленных (в конечных позициях) элементов, что дало возможность говорить о так называемой «внетекстовой событийной оси» произведений. То есть в самом тексте описана как будто лишь часть происходящих событий, на наличие полной событийной картины указывают сильные позиции текста;

3) наличие большого числа повторяющихся микротем разных регистров, следующих друг за другом в неизменном порядке. Наличие подобной линейной симметрии наблюдалось в большинстве текстов на уровне микротем и даже регистровых блоков, что позволяет говорить об особом нарративном ритме жанра немецкого короткого рассказа.

Краткое ознакомление с текстами других жанров малых форм показывает, что выделенные для жанра короткого рассказа особенности могут присутствовать и в текстах других жанров, но не в совокупности, а по отдельности. Это дает основание говорить о существовании именно жанровых особенностей текстов коротких рассказов на логико-семантическом уровне. Однако для получения более достоверных результатов необходим выход за пределы жанра с использованием той же методики и верификации полученных данных на большем объеме фактического материала.

Т. В. Бусел

МОБИЛЬНЫЕ ПРИЛОЖЕНИЯ КАК ИНСТРУМЕНТ ФОРМИРОВАНИЯ МОТИВАЦИИ СТУДЕНТОВ К ИЗУЧЕНИЮ ИНОСТРАННОГО ЯЗЫКА

Прошедшие месяцы отмечены появлением большого количества научных работ, посвященных проблемам мобильного обучения. В них рассматриваются понятие и типы мобильного обучения, проблемы создания образовательных мобильных приложений, перспективы внедрения мобильных технологий в образование и т.п. Педагогические сообщества активно обсуждают внедрение мобильных технологий в практику преподавания, анализируют их достоинства и недостатки, делятся опытом на специальных форумах по использованию мобильных технологий.

Мобильные устройства (*smart mobile devices*) успешно используются при изучении различных учебных дисциплин, и иностранный язык не является исключением. Мобильные технологии в образовании базируются в основном на использовании мобильных телефонов, которые в силу постоянного совершенствования могут рассматриваться не только как средство связи, но и как средство обучения.