

ТЕОРИЯ СЕГМЕНТНОЙ РЕПРЕЗЕНТАЦИИ ДИСКУРСА ДЛЯ РЕШЕНИЯ ЗАДАЧ СУДЕБНОЙ ЛИНГВИСТИЧЕСКОЙ ЭКСПЕРТИЗЫ ПРИ ИЗВЛЕЧЕНИИ ИЗ ТЕКСТА ИМПЛИЦИТНОЙ ИНФОРМАЦИИ

SEGMENTED DISCOURSE REPRESENTATION THEORY FOR SOLUTION OF FORENSIC LINGUISTICS PROBLEMS IN IMPLICIT INFORMATION EXTRACTION

Д.С. Кондрашова (dkond@rambler.ru)

Московский государственный университет им. М.В. Ломоносова

Семантическая экспертиза, как один из видов судебной лингвистической экспертизы (СЛЭ), нацелена на выявление смыслов в тексте и их анализ с разных точек зрения. Для решения задач такой СЛЭ предлагается использовать **теорию сегментной репрезентации дискурса** [Asher, Lascarides 2003].

Перефразируя Е.В. Падучеву [Падучева 1976: 91], можно утверждать, что достижения в области семантики существенно расширяют доступные лингвистике сферы анализа языка. Одной из таких сфер стала разрабатываемая в настоящее время – в рамках судебной лингвистической экспертизы (СЛЭ) – **семантическая экспертиза**.

Практика проведения СЛЭ получила в последнее время довольно широкое распространение. К ведению СЛЭ относится решение широкого круга прикладных задач – идентификационных, классификационных и диагностических, – возникающих в связи с анализом речевой деятельности. Объектом СЛЭ являются продукты речевой деятельности (от отдельного слова до целого текста или группы текстов), зафиксированные в письменной форме. Продукт речевой деятельности может оцениваться с точки зрения цели его порождения, выражаемого им смысла, способа выражения этого смысла и характера его воздействия на адресата или аудиторию. СЛЭ опирается на разные виды лингвистического анализа, такие как семантический, грамматический, орфографический и др.

Сегодня можно говорить о трех основных типах задач, решаемых в рамках СЛЭ:

- (1) анализ особенностей текста как носителя информации о его авторе (автороведческие исследования);
- (2) анализ текста с целью выявления выраженных в нем смыслов и характера воздействия текста на адресата или аудиторию (смысловые, или семантические, исследования);
- (3) анализ наименований (названий, торговых марок, доменных имен и т.д.) как средства обозначения «продуктов» человеческой деятельности (ср. [Возможности производства судебной экспертизы... 2004: 421-422]).

Потребность в лингвистических экспертизах все время возрастает, что делает разработку их методик актуальнейшей задачей. В особенности это касается одной из разновидностей СЛЭ, **семантических экспертиз** (например, связанных с делами по защите чести и достоинства), ставших предметом внимания настоящего исследования.

Семантическая экспертиза главным образом нацелена на выявление смыслов, передаваемых исследуемыми текстами, и анализ этих смыслов с разных точек зрения.

Совершенно очевидно, что для решения задач такой экспертизы необходимо создание особой экспертной методики, позволяющей выявлять не только эксплицитно выраженные смыслы, но и скрытые, имплицитные смыслы, поскольку любое ЕЯ высказывание содержит значительно больше информации, чем выражается в нем в явной форме. Одновременно встает вопрос об увеличении степени объективности семантической СЛЭ, т.к. смысловой анализ текста, по мнению многих ученых, является скорее искусством, чем формальной процедурой. На наш взгляд, существуют такие теоретические разработки в области пересечения современной лингвистической семантики и лингвистической прагматики, которые могут помочь решить оба вышеупомянутых вопроса. Мы предлагаем использовать с этой целью **теорию сегментной репрезентации дискурса** (Segmented Discourse Representation Theory, SDRT) [Asher, Lascarides 2003], отличающуюся от классической версии (DRT, [Kamp, Reyle 1993]), в частности, учетом риторических связей.

В рамках поставленных проблем центром нашего внимания является имплицитный пласт содержания подвергающихся экспертному исследованию текстов. Базовое понятие имплицитного рассматривается нами как языковой уровень выражения смысла¹. Понятие же имплицитной информации мы понимаем следующим образом.

«Имплицитная информация представляет собой *неоднородную* по своему составу часть совокупной информации, передаваемой произвольным текстом естественного языка, *не содержащуюся в тексте в явном виде, но сосуществующую* с непосредственно выраженной (эксплицированной лексико-семантической структурой текста) информацией, *извлекающуюся или могущую быть извлеченной* из данного текста – *вне зависимости от* ее характера, актуальности для того или иного участника общения и ее «запланированности» адресантом, *а также от* того, проявляется ли эта информация в изолированном рассматриваемом тексте или только при включении данного текста в более широкий контекст (ср. [Долинин 1983: 37], а также [Падучева 1981: 23] и [Кобозева 2000: 215])» [Кондрашова 2005]. О разнообразии видов имплицитной информации и о принимаемом нами статусе пресуппозиции в этом аспекте можно также прочесть в [Кондрашова 2005].

В исследовании, которому посвящена настоящая статья, рассматриваются примеры анализа случаев наличия в тексте той или иной информации в имплицитной форме при проведении семантической экспертизы. На основании собранных данных был составлен обзор средств, которые лингвисты-эксперты используют для описания таких случаев. В ходе критического осмысления полученных результатов возник вопрос о необходимости разработки единой терминологической системы для анализа скрытой информации в ходе экспертного исследования, а также о необходимости создания самой методики анализа и повышения его объективности.

Материалом для исследования послужили тексты реальных судебных семантических экспертиз (т.е. тексты экспертных исследований, составленные экспертами) в сопровождении с соответствующими подвергнутыми экспертизе текстами (т.е. текстами, которые анализировали эксперты при проведении этих экспертиз).

Данные имеющихся в нашем распоряжении экспертиз позволяют утверждать, что анализ имплицитного плана содержания исследуемых текстов проводится в опоре, главным образом, на интуицию исследователя. Применение же элементов SDRT позволяет привести процесс анализа к некоторому подобию алгоритма, повысив степень объективности семантической экспертизы.

SDRT является динамической семантической теорией интерпретации дискурса и представляет собой попытку моделирования семантико-прагматического интерфейса: особое внимание уделяется тому, как процесс интерпретации высказывания включает в себя обогащение его композициональной и лексической семантики дополнительным содержанием. Это позволяет объяснить несколько взаимосвязанных явлений дискурса, в частности:

- разрешение анафоры различных типов (включая пресуппозиции);
- разрешение лексической полисемии;
- импликатуры дискурса.

Поскольку в рамках DRT, которая лежит в основе SDRT, эти проблемы не могут быть решены, авторам пришлось расширить классическую теорию².

Охарактеризовать концепцию [Asher, Lascarides 1998] (где впервые используется новый подход к пресуппозициям) можно при помощи двух ее основных черт:

1) авторы принимают вслед за [Van der Sandt 1992] анафорическую природу пресуппозиций³, но дают анафорической теории новый поворот: предполагается, что пресуппозиции, как и другие анафорические выражения, например, анафорические местоимения, имеют *недоопределенное* (underspecified) семантическое содержание⁴. Интерпретация пресуппозиций в контексте включает разрешение недоопределенности.

2) вторая отличительная черта касается способа разрешения недоопределенности. Предлагается формальная модель семантики дискурса (SDRT), в которой семантическая недоопределенность в пропозиции разрешается определением способа, которым эта пропозиция риторически связана с контекстом. Тот факт, что выбор риторической связи оказывает влияние на семантическое содержание, является центральной идеей этой теории.

Принятие таких установок имеет три следствия:

¹ Уровневая концепция понимания имплицитного описывается в [Кондрашова 2005].

² SDRT представляет из себя расширенный вариант DRT с обогащенным понятием структуры дискурса и видоизмененной функцией обновления дискурса [Asher 1993]; [Lascarides & Asher 1993].

³ В общих чертах эта стратегия интерпретации пресуппозиций описана в работе [Выборнова 2002].

⁴ Понятие семантической недоопределенности (underspecification) подробно описывается в работе [Asher, Lascarides 1993] в соответствии с концепцией [Reyle 1993].

а) первое затрагивает прагматику. Поскольку прагматическая информация играет роль в вычислении риторических связей, она также ограничивает и интерпретацию пресуппозиций. SDRT позволяют формально эксплицировать способ взаимодействия семантики и прагматики при вычислении риторической связи между пропозицией и контекстом в дискурсе. Это взаимодействие и определяет интерпретацию пресуппозиций и, таким образом, дает более богатый источник ограничений на пресуппозиции, чем стандартные подходы.

б) второе следствие интерпретации пресуппозиций в SDRT касается приспособления. Во всех предшествующих динамических теориях пресуппозиции приспособление представляет собой просто добавление, а не соотнесение пресуппонируемого содержания с некоторой доступной частью контекста. Такой механизм в этих теориях используется исключительно для пресуппозиций, он не применяется для объяснения каких-либо других языковых явлений, в том числе и анафорических. Тогда как [Asher, Lascarides 1998] моделируют пресуппозиции исключительно в терминах процедуры обновления дискурса, используемой в SDRT. Понятие добавления пресуппозиций к контексту заменяется понятием их риторической связи с ним: привязывание пресуппозиций к контексту при помощи риторических связей заменяет их приспособление. Принимая во внимание, что теория риторической структуры в SDRT используется для моделирования широкого ряда языковых явлений в применении к асертивной информации (в частности, для разрешения семантической неоднозначности), было бы странно, если пресуппонируемая информация была абсолютно нечувствительна к риторической функции. Такая концепция пресуппозиций позволяет интегрировать их вычисление в стандартную процедуру обновления дискурса (*discourse update procedure*), заменяющую простое добавление и предусмотренную любой теорией представления дискурса в духе DRT, что унифицирует и, следовательно, упрощает весь процесс репрезентации и интерпретации дискурса; кроме того, авторы [Asher, Lascarides 1998] утверждают, что главной чертой их модификации процедуры обновления дискурса является то, что она «моделирует, каким образом содержание дискурса представляет из себя нечто большее, чем то, что эксплицитно выражено языковыми средствами» (!) [Asher, Lascarides 1998: 265].

в) третье следствие, вытекающее из второго, предполагает, возможность композиционного понимания пресуппозиции. В SDRT значение целого текста является производным от значения его частей и способа их соотнесения друг с другом. Пресуппозиции, как и асертивные компоненты содержания, порождают недоопределенные, но «интерпретабельные» логические формы. Процедура построения семантической репрезентации дискурса разрешает часть недоопределенных элементов и соотносит их друг с другом при помощи дискурсивных отношений, представляющих их риторическую функцию в дискурсе. Таким образом, пресуппозиции могут делать свой вклад в общее содержание дискурса. В этом смысле их природа не менее композициональна, чем у асертивных компонентов содержания.

Данный подход предлагает более широкий взгляд на пресуппозицию. Существующие в настоящее время динамические объяснения пресуппозиций сосредоточены на феномене проблемы проекции (*Projection Problem*). В понимании [Asher, Lascarides 1998] проблема проекции пресуппозиций представляет собой имеющий большое значение и применяемый к простым предложениям в дискурсе частный случай более общей проблемы: каким образом пресуппозиции, порождаемые элементами полисентенциального дискурса, влияют на его структуру и содержание?

Для формальной репрезентации текста в SDRT используются структуры представления дискурса (SDRSs), включающие структуры для репрезентации содержания клауз (DRSs) и риторические отношения типа *Parallel*, *Narration* и *Background* для репрезентации риторических функций DRSs в контексте, причем эти отношения могут оказывать влияние на содержание клауз, которые они соотносят и, следовательно, на содержание дискурса в целом.

Репрезентация дискурса происходит путем последовательной интерпретации каждой порции новой информации, которая рассматривается при этом как вносящая изменения в текущий контекст, т.е. как имеющая определенный потенциал изменения контекста (*context change potential*). Сначала при репрезентации используется грамматика для композиционного построения DRSs, а затем уже из них, в соответствии с формально определенной процедурой обновления дискурса, динамически строится SDRS. Для обновления текущей SDRS применяется связующая логика (*glue logic*), при помощи которой определяется место присоединения новой DRS и выводится риторическое отношение для этого присоединения (эта последняя функция связующей логики базируется на выводах, лежащих в основе естественного рассуждения (*commonsense entailment*)).

Рассмотрим следующий фрагмент дискурса:

- (1) а. Проблема снабжения в области была решаемой;
- б. технику перевезли из Калуги в Обнинск;
- в. Чуйко отобрал два комбайна
- д. и переправил их в совхоз.

Грамматика даст нам две структуры для (1a): одну для репрезентации пресуппонируемого содержания (*в области имеет место проблема со снабжением*), а другую для репрезентации асертивного содержания (*она решается*). Поскольку предложение является первым в дискурсе, SDRS для всего предложения должна быть обновлена сначала асертивным содержанием (добавлением его к пустому множеству – нулевому контексту начала дискурса), так как пресуппонируемое содержание требует (риторической) анафорической привязки. Теперь пресуппонируемое содержание должно быть присоединено к введенному асертивному содержанию при помощи риторического отношения. Содержание и пресуппозиции и асерции стативно (в обоих случаях перед нами состояние), поэтому согласно связующей логике для их связи мы применяем отношение *Background*, т.е. тот факт, что в области имеет место проблема со снабжением, является фоновой информацией по отношению к тому, что эта проблема решается.

Не имея возможности подробно описать процесс репрезентации всего примера, мы опустим вывод риторического отношения *Explanation*, присоединяющего к (1a) последующий дискурс и интерпретацию последней части дискурса (1d), и сделаем акцент на том, каким образом формальная репрезентация может обогащаться содержанием, не имеющим эксплицитного выражения в поверхностной структуре текста.

Определенная ИГ *два комбайна* в (1c) означает, что DRS, репрезентирующая пресуппонируемое содержание (1c), содержит недоопределенное отношение соотнесения (bridging relation) $B=?$ с некоторым недоопределенным предшествующим объектом $v=?$ (т.е. раз ИГ определенная (выражает известные, по мнению говорящего, референты), то что именно это были за комбайны?)⁵. Согласно правилам связующей логики, асертивное содержание (1c) соединяется с содержанием (1b) при помощи отношения *Elaboration*, которое, как и *Narration*, налагает на соединяемые структуры пространственно-временные ограничения. И мы получаем, что *два комбайна* являются теми комбайнами, которые находятся в Обнинске (и которые, соответственно, были привезены вместе со всей техникой из Калуги). Благодаря устройству процедуры обновления дискурса, именно эта дополнительная информация, добавляемая к эксплицитному содержанию DRS, разрешает $B=?$ и $v=?$ в $B=$ «(находиться) в» и $v=$ «Обнинск», соответственно.

Таким образом, этот пример демонстрирует, что разрешение недоопределенности в пресуппозиции определяется способом, которым асертивное содержание присоединяется к контексту. Возможно и обратное: способ присоединения пресуппозиции к контексту может разрешать недоопределенные условия в DRS, репрезентирующей асертивное содержание.

На первый взгляд, данный пример может показаться слишком простым для обоснования привлечения сложной теории репрезентации дискурса, однако, как показывает собранный нами материал, в запросах к экспертам могут фигурировать намного более сложные случаи, в которых интуиция и языковое чутье исследователя могут оказаться недостаточными для убедительных доказательств. Более того, и в таких простых случаях может быть назначена экспертиза. Также мы полагаем, что применение единой методики для анализа произвольного ЕЯ текста может существенно повысить степень объективности семантической экспертизы.

Важно также отметить, что необходимо учитывать специфические особенности языка экспертов, которые сформировались под влиянием использования результатов экспертных исследований в судебной практике в связи с существенным различием багажа лингвистических знаний у экспертов-лингвистов и у служащих суда: язык экспертизы должен быть понятен последним. В таких изданиях как [Цена Слова 2002] и [Памятка 2004] можно найти довольно ограниченные по своему составу списки слов, выражающих лингвистические понятия, которые, как правило, используются при фиксации процесса проведения экспертного исследования и записи его результатов (заключения экспертов). Списки сопровождаются краткими и упрощенными толкованиями, предназначенными для служащих суда. Таким образом, традиционная лингвистическая терминология оказывается ограниченной, упрощенной и не всегда достаточной даже для описания типичных и обычно не вызывающих затруднения случаев. Более того, в рамках экспертной деятельности для многих явлений не существует единого подхода к их описанию.

Возможно, впоследствии произойдет пересмотр концепции языка семантической экспертизы в связи с нетривиальными задачами, решаемыми при ее проведении, а также с целью приведения к единообразию формата исследования и, отчасти, повышения степени объективности экспертиз⁶.

Согласно сказанному, а также в связи с тем, что применение SDRT на практике требует достаточно глубокого уровня математических и формально-логических познаний, которыми владеет далеко не каждый лингвист-эксперт, на данном этапе определенное упрощение предлагаемого в описываемом исследовании подхода к решению проблем семантической экспертизы является необходимым. Формальная же модель SDRT

⁵ По данным материала исследования, лингвистам-экспертам очень часто приходится устанавливать и доказывать наличие такого рода соотношения в исследуемом тексте.

⁶ Для сравнения, можно упомянуть о существовании таких экспертиз как химические, медицинские и проч., не предполагающие упрощенного понимания соответствующей терминологии. В случае семантической экспертизы некоторое упрощение допустимо, однако необходимо единообразие и системность в ее использовании.

может использоваться для расширения и обогащения данных современной семантической теории и служить обоснованием использования упрощенной версии модели в практике проведения семантических экспертиз.

Отсюда вытекает следующий формат анализа материала в данном исследовании:

1. релевантный отрывок из исследуемого экспертизой текста;
2. вопрос, поставленный перед экспертами;
3. выдержки из экспертного исследования, посвященные приведенному выше отрывку;
4. анализ данного отрывка с современной лингвистической точки зрения;
5. анализ данного отрывка с позиций формальной модели SDRT;
6. адаптированный для применения в экспертном исследовании анализ данного отрывка.

В заключение хочется еще раз подчеркнуть, что разработка методики извлечения из текста имплицитной информации особенно важна для целей проведения судебных лингвистических исследований, ведь, например, такой имплицитный компонент как пресуппозиция «является той базой, которая дает возможность формировать мысли и выражать их различными языковыми средствами» [Выборнова 2002: 63], а, следовательно, и определенным образом влиять на сознание адресата.

Мы надеемся, что настоящее исследование может внести вклад как в практическое решение обсуждаемых в статье задач судебной семантической экспертизы, так и в теоретическое осмысление семантики естественного языка, в частности, касающееся имплицитного плана содержания дискурса.

Список литературы

- Asher, N. & Lascardes, A. *The Semantics and Pragmatics of Presupposition*. // *Journal of Semantics*. 1998. Vol. 15. № 3, pp.239-299.
- Asher, N. & Lascardes, A. *Logics of conversation*. Cambridge etc.: Cambridge university press, 2003.
- Kamp, H. & Reyle, U. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer Academic Publishers, 1993.
- Reyle, U. *Dealing with Ambiguities by Underspecification: Construction, Representation and Deduction*. // *Journal of Semantics*. 1993. Vol. 10. № 2, pp.123-179.
- Van der Sandt, R.A. *Presupposition projection as anaphora resolution*. // *Journal of Semantics*. 1992. Vol.9. № 4, pp.333-377.
- Возможности производства судебной экспертизы в государственных судебно-экспертных учреждениях Минюста России. / Под общ. ред. начальника Управления судебно-экспертных учреждений Министерства юстиции РФ, к.х.н. Т.П. Москвиной. // М.: ИПК РФЦЭС, 2004.*
- Выборнова О.В. *Пресуппозиционный компонент общения и его прикладное моделирование*. // *Труды международного семинара Диалог'2002 по компьютерной лингвистике и ее приложениям*. М.: 2002. Т. 1. Теоретические проблемы. С.63-67.
- Долинин К.А. *Имплицитное содержание высказывания*. // *Вопросы языкознания*. 1983. № 6. С. 37-47.
- Кобозева И.М. *Лингвистическая семантика*. // М.: Эдиториал УРСС, 2000
- Кондрашова Д.С. *К проблеме классификации типов имплицитной информации*. // *Обработка текста и когнитивные технологии: Сборник статей*. № 11. (Материалы VIII Международной конференции 'Cognitive Modelling in Linguistics', Varna, 2005). / Под ред. В. Соловьева и др. Казань: Казанский государственный университет, 2006. Т. 1. С. 245-252.
- Падучева Е.В. *Понятие презумпции в лингвистической семантике*. // *Семиотика и информатика*. М.: ВИНТИ, 1976. Вып. 8. С. 91-124.
- Падучева Е.В. *Презумпции и другие виды неэсплицитной информации*. // *Научно-техническая информация*. 1981. Сер.2. № 11. С. 23-30.
- Памятка по вопросам назначения судебной лингвистической экспертизы: Для судей, следователей, дознавателей, прокуроров, экспертов, адвокатов и юристконсультантов. / Под ред. проф. М.В. Горбаневского. // М.: «Медведь», 2004.*
- Цена слова: Из практики лингвистических экспертиз текстов СМИ в судебных процессах по защите чести, достоинства и деловой репутации. / Под ред. проф. М.В. Горбаневского – 3-е изд., испр. и доп. // М.: «Галерея», 2002.*