

Теория лингвистических экспертиз как направление прикладной лингвистики

А. Н. Баранов

В докладе обсуждается теория лингвистических экспертиз. Последняя рассматривается в как направление прикладной лингвистики и как составная часть «судебной лингвистики». Лингвистические экспертизы предполагают использование знаний о структуре и функционировании языка, полученные в рамках теоретической лингвистики. Предлагаемая типология лингвистических экспертиз включает официальные и неофициальные экспертизы, экспертизы звучащей речи, экспертизы письменного текста, экспертизы товарных знаков, автороведческие экспертизы, вербально-визуальные экспертизы. Более подробно обсуждаются автороведческие экспертизы. Предлагается новый метод установления авторства текста – **лингво-статистический анализ неполнозначной лексики**.

1. Введение в проблему

Демократизация общества (как бы она ни понималась и ни оценивалась) и возникновение сферы публичной политики с начала 90-х гг. XX в. привели к появлению новых дисциплин прикладной лингвистики. К ним можно отнести возникновение на отечественной почве политической лингвистики, формирование направления изучения языка рекламы, развитие прикладных методик лингвистического мониторинга языка средств массовой информации – как для целей лингвистической теории и культуры речи, так и для выявления изменений в структуре общественного сознания по данным языка и т.д. К таким успешно развивающимся направлениям прикладного языкознания относится и **лингвистическое консультирование** – использование знаний о структуре и функционировании языка для оптимизации использования языка в СМИ, публичной политике, в преподавании, в законодательной деятельности, в юридической сфере, в обеспечении эффективного документооборота и т.д. Аналогом данного направления в области политологии можно считать **политическое консультирование**, целью которого является применение знаний о способах взаимодействия с общественностью, инструментах исследования общественного сознания, о практике проведения избирательных кампаний для организации деятельности политических субъектов различного уровня (см. по этому поводу, например, [Политическое консультирование 1999]). В отличие от политического консультирования, в область лингвистического консультирования непосредственно входит обеспечение проведения лингвистических экспертиз различной целевой направленности – и прежде всего для судебной сферы, области патентования и товарных знаков. Лингвистические экспертизы в судебной сфере иногда называются в литературе **юрислингвистикой** (см., например, [Юрислингвистика 2002]). В Германии близкие направления имеют наименования - *forensische Linguistik* букв. «судебная

лингвистика», Rechtslinguistik букв. «правовая лингвистика» [Untersuchungen 1986]. При этом forensische Linguistik чаще всего понимается как использование знаний о функционировании языка в проведении следствия, а Rechtslinguistik охватывает область функционирования языка в суде, языковую форму юридической аргументации, языковые требования к оформлению юридических документов, интерпретацию семантики текстов документов и т.д. Очевидно, что сфера лингвистических экспертиз, связанных с использованием знаний о функционировании языка, не совпадает ни с «юрислингвистикой», ни с «правовой лингвистикой». Например, лингвистическая экспертиза товарного знака не обязательно связана с юриспруденцией, хотя и может использоваться для регулирования правовых отношений. Аналогично, автороведческая экспертиза – экспертиза по установлению авторства текста – не всегда предусматривает судебное разбирательство. Ср. в этой связи многочисленные исследования авторства «Тихого Дона», произведений У. Шекспира, М.Е. Салтыкова-Щедрина, В.В. Маяковского и др., имеющие, скорее, общественный резонанс, чем судебную перспективу. Сам термин «экспертиза» понимается в этом случае несколько расширительно – как выражение мнения профессионального лингвиста о тех или иных языковых феноменах. При этом существенно, что это мнение используется за пределами лингвистики – в последнем случае в литературоведении, истории и т.д. В силу сказанного для обозначения данной области лингвистического консультирования естественнее использовать более расширительный термин, например, «лингвистическое экспертоведение», «лингвистическая экспертология» или «теория лингвистических экспертиз». Первые два термина содержат неологизмы, степень приемлемости которых неочевидна. В дальнейшем в качестве рабочего термина в указанном значении используется выражение **лингвистическая экспертология** или **теория лингвистических экспертиз**.

2. Типология лингвистических экспертиз

Классификация лингвистических экспертиз может проводиться по различным параметрам. Формальные критерии не представляют существенного интереса с точки зрения собственно теории лингвистических экспертиз, хотя и влияют на процедуру выполнения экспертизы и способ представления результата. По формальному параметру лингвистические экспертизы (и не только собственно лингвистические) разделяются на две большие группы: официальные и инициативные. **Официальные** экспертизы выполняются по постановлению суда или органов дознания и имеют статус доказательств. **Инициативные** экспертизы выполняются по инициативе любых заинтересованных физических и юридических лиц (в том числе ответчиков и истцов, адвокатов, частных фирм, государственных организаций, патентных поверенных и т.д.). Инициативные экспертизы могут признаваться как правомерное доказательство только по решению суда или органов дознания. По количеству экспертов различаются комиссионные и неkomиссионные экспертизы. **Комиссионные экспертизы** выполняются несколькими экспертами, а неkomиссионные – только одним. По объекту лингвистических экспертиз выделяются экспертизы звучащей речи, экспертизы письменного текста и вербально-визуальные экспертизы. В **экспертизах звучащей речи** исследуются акустические характеристики речевых высказываний, в частности, для определения автора речевого сообщения. Если в экспертизах звучащей речи используется специальный инструментарий акустики, фонетики и фонологии, то **экспертизы письменного текста** разнообразны по своей направленности и используемым методикам. В экспертизах такого рода может использоваться и морфологический анализ (например, при исследовании сходства и различий товарных знаков), и синтаксический анализ (например, при установлении синтаксической сложности текстов при установлении авторства), и семантический анализ (например, семантическое исследование инвектив в экспертизах по

делам о защите чести и достоинства), и анализ текста (например, в экспертизах по определению семантического подобия текстов), и лингво-статистический анализ (например, в экспертизах по установлению авторства).

Вербально-визуальные экспертизы в качестве объекта для исследования имеют сложно организованный феномен – сочетание вербальной и невербальной информации, то есть комбинация текста и изображения. Это оказывается необходимым, например, при оценке степени близости товарных знаков, когда по набору символов, графем или «буквоподобных» элементов товарные знаки близки, но изобразительно поданы по-разному (ср. дело по товарным знакам **AIGLE** и **EGLE**¹). Анализ изображения и текста в единстве требуется и в тех случаях, когда семантика текста или его фрагментов, иллюстрирующих изображение, не вполне ясна. Так, выражение *Знай наших!* может использоваться и как фразеологизм в значении ‘выражение говорящим положительных эмоций, вызванных чем-либо, интерпретируемым как знак превосходства лица или группы лиц, с которыми говорящий себя отождествляет’, и как свободное словосочетание – императив-рекомендация, коммуникативная цель которого заключается в том, чтобы рекомендовать адресату пополнить свои знания о конкретной группе лиц, с которыми говорящий себя отождествляет. В одном из экспертных дел предметом исследования оказалась фраза *Знай наших!*, помещенная на банке напитка «Coca-Cola»². Претензия заключалась в том, что эта фраза была ранее закреплена как слоган в составе товарного знака другого производителя, причем именно в функции фразеологизма. Анализ фрагментов текста и изображения на банке «Coca-Cola» позволил сделать вывод о том, что в рассматриваемом случае фраза *Знай наших!* использована в прямом значении и, следовательно, не может рассматриваться как заимствование слогана из товарного знака другого производителя.

В ряде случаев вербально-визуальная экспертиза требует анализа целого видеоряда в его сочетании с вербальными элементами. Так, исследование видеоряда, его текстовых составляющих и текста «за кадром» позволило показать, что мини-сюжет о производстве парфюмерной продукции фирмы «МИРРА» и иллюстрирующий ее «закадровый» текст, с одной стороны, и фраза о распространении фальшивой парфюмерии и косметики из мини-сюжета о проверке вьетнамского рынка – с другой, связаны отношением экземплификации. То есть фирма «МИРРА» связывается в телевизионной передаче с производством нелегальной продукции.

Еще один тип экспертиз – **автороведческие** экспертизы, задачей которых является установление авторства текста. Остановимся на этом типе экспертиз несколько подробнее.

3. Автороведческие экспертизы

Типовые случаи автороведческих экспертиз описываются следующими ситуациями:

- **А. Множественная неопределенность.** Имеется множество текстов или их фрагментов. Необходимо установить, сколько авторов их писали и как распределен авторский вклад по каждому конкретному тексту.
- **Б. Сравнение по образцу.** Имеется пример текста (текстов) некоторого автора *X*. Необходимо установить, является ли он и автором некоторого другого текста (текстов).
- **В. Конкуренция образцов.** Имеются образцы текстов авторов *X, Y, Z...* Необходимо установить, кто из них является автором текстов T_1, T_2, \dots, T_n .

¹ Экспертиза, проведенная автором, совместно с Ю.К. Пироговой и Э.И. Почтарь.

² Экспертиза проводилась автором совместно с О.М. Грунченко.

Ситуация «Множественной неопределенности» представляет собой наиболее сложный случай установления авторства. Поскольку нет эталонных корпусов текстов предполагаемых авторов, то все имеющиеся тексты должны быть проверены на однородность с точки зрения авторского стиля. Если не исключаются случаи соавторства (характер которого, вообще говоря, может сильно варьировать), то при современных методиках установления авторства получение какого-то разумного результата в такой ситуации вряд ли возможно. Впрочем, реально случаи такого рода практически не встречаются.

Наличие гипотезы, формулируемой на экстралингвистических основаниях, существенно упрощает установление авторства. Именно такой случай представлен ситуациями «Сравнения по образцу» и «Конкуренции образцов».

Современные формальные методики установления авторства текста (именно такие методики установления авторства в первую очередь используются в экспертизах) опираются на стохастическую модель порождения речи. В основе этой модели лежит представление о том, что с возрастанием объема текста авторские особенности языкового оформления содержания становятся устойчивыми с вероятностной точки зрения, что позволяет устанавливать авторство по стабильно повторяющимся формальным характеристикам текста. В стилиметрии – относительно новом междисциплинарном научном направлении, возникшем в 60-е гг. XX в. [Семиотика... 1972] – и в автороведческих исследованиях широко используется метод **статистического анализа формальных характеристик текста** (см., например, [Ворончак 1972; Батов, Сорокин 1975]). По сути, он сводится к выявлению значений таких параметров, как объем словаря, количество полных и служебных слов, объем индивидуального словаря, количество предложений в тексте, средняя длина предложения, отношение числа глаголов к общему количеству словоупотреблений в тексте, отношение числа предлогов к общему количеству словоупотреблений в тексте и т.д. Предполагается, что авторский стиль отражается в уникальном наборе значений таких формальных характеристик. Достигнутые с помощью этого метода результаты обнадеживают, хотя и оставляют целый ряд вопросов нерешенными. В частности, принципиальный вопрос о границах нормы авторского варьирования формальных характеристик остается без ответа. Кроме того, многие формальные параметры с лингвистической точки зрения вряд ли имеют отношение к авторству. Например, соотношение между количеством личных местоимений и общим количеством словоупотреблений; часто используемая оценка удельного веса названий по отношению к авторскому словарю в целом и др. трудно интерпретировать как характеристики стиля автора.

В последнее время методика статистического анализа формальных характеристик текста подвергалась различным модификациям. Так, перспективный вариант этой методологии, основанный на векторном представлении значений формальных параметров стиля, предложен М.А. Марусенко [Марусенко 1996]. Иной подход, хотя и косвенно связанный с методикой формальных характеристик, предложен автором в [Баранов 2000] и ряде автороведческих экспертиз. Суть этого метода заключается в следующем. В рамках относительно небольшого текста в 1000-2000 словоупотреблений сочетание значений формальных характеристик – аналогов авторского стиля – не получает адекватного отображения. Кроме того, не проявляются в должной мере и многие другие характеристики, например, особенности использования авторской фразеологии и идиоматики, а также метафорической системы, системы эпитетов и т.д.

Иными словами, использование методов установления авторства, основанных на анализе формальных характеристик, или, наоборот, на чисто стилистическом или литературном анализе текста, не представляется возможным. С другой стороны, грамматические особенности авторского стиля – частота употребления неполнозначных, "строевых" слов (частиц, союзов, предлогов, некоторых модальных слов, вводных выражений) – для текстов

порядка 1 000 - 2 000 словоупотреблений сохраняются. Это и дает возможность подтверждать или опровергать имеющиеся гипотезы относительно авторства текста. Подчеркнем, что гипотеза об авторстве в любом случае должна присутствовать, причем она должна формулироваться на экстралингвистических основаниях. Такой метод установления авторства текста может быть назван **лингво-статистическим анализом неполнозначной лексики**.

Данный метод предполагает наличие двух корпусов текстов – эталонных текстов, представляющих авторский стиль предполагаемого автора (или авторов), и анализируемых (спорных) текстов. Далее анализ проходит в виде эксперимента, в ходе которого на различных этапах осуществляются следующие процедуры:

- составление словарей эталонного корпуса (корпусов – если предполагаемых авторов несколько) и корпуса (корпусов) спорных текстов с указанием абсолютной и относительной частоты употреблений в соответствующем корпусе;
- удаление полнозначных лексем из полученных словарей;
- сопоставление словарей, включающих только неполнозначную лексику, для выявления лексем с близкой частотой; порог близости частоты определяется в каждом конкретном случае по специальной методике;
- формирование кластеров «авторских» лексем, включающих только те неполнозначные слова, которые являются близкими по относительной частоте для данной пары сравниваемых корпусов.

В нескольких проведенных экспериментах было установлено, что публицистические тексты, принадлежащие одному и тому же автору, содержат от 60 до 90 общих «авторских» лексем. Кластеры авторов, очевидно различающихся по стилю, содержат от 20 до 35 «авторских» лексем.

Развитие теории лингвистических экспертиз и как теоретического и как прикладного направления предполагает решение ряда важных задач. Прежде всего необходимо повысить степень объективности результатов лингвистических экспертиз для того, чтобы экспертизы были воспроизводимы другими исследователями. Разумеется, это возможно только при использовании достаточно формализованных процедур лингвистического анализа. Иными словами, проблема методов лингвистического исследования, во многом забытая в современной теоретической лингвистике, встает в полный рост. Второе направление деятельности в лингвистической «экспертологии» - перевод терминологии законодательства в лингвистические термины и наоборот. Многие понятия из сферы лингвистических экспертиз по защите чести и достоинства до сих пор не получили адекватной лингвистической интерпретации, хотя некоторое продвижение в этой области имеется, ср., например, [Баранов и др. 2002].

Наконец, третье весьма перспективное направление – исследование закономерностей организации вербальной и визуальной составляющих текста в семиотическом понимании. Современная массовая культура с очевидностью движется в сторону визуализации информации. На этом во многом основаны такие феномены, как электронные СМИ, реклама, гипертекстовые технологии организации знания. Дискурс оказывается не только вербальным, но и визуальным. Очевидно, что доля речевых образцов, прямо связанных с изображением, будет только возрастать. Понятно, что будет возрастать и доля соответствующих экспертиз. Между тем лингвистическая теория пока не дает надежных инструментов исследования таких рече-визуальных феноменов.

Литература

1. Баранов А.Н. Введение в прикладную лингвистику. М.: УРСС, 2000.

2. Баранов А.Н., Бельчиков Ю.А., Сафонова Ю.К., Шварцкопф Б.С. Цена слова. Из практики лингвистических экспертиз текстов СМИ в судебных процессах по защите чести, достоинства и деловой репутации. М.: Галерея, 2002, 424 с. Изд. 3-е издание, исправленное и дополненное.
3. Батов В.И., Сорокин Ю.А. Атрибуция тексте на основе формальных характеристик // Изв. АН СССР, сер. лит. и языка, 1975, т. 34, № 1.
4. Ворончак Е. Методы вычисления показателей лексического богатства текстов // Семиотика и искусствометрия. М., 1972.
5. Марусенко М.А. Атрибуция анонимных и псевдонимных текстов методами прикладной лингвистики // Прикладное языкознание. СПб, 1996.
6. Политическое консультирование / под ред. Е.В. Егоровой-Гантман. М., 1999.
7. Семиотика и искусствометрия. М., 1972.
8. Юрислингвистика-3: проблемы юрислингвистической экспертизы: Межвуз. сб. науч. тр./ Под ред. Н.Д. Голева. Барнаул, АГУ, 2002
9. Untersuchungen zur Rechtslinguistik. Interdisziplinäre Studien zu praktischer Semantik und Strukturierender Rechtslehre in Grundlagen der juristischen Methodik. / Müller, Friedrich (Hrsg.). Berlin: Duncker & Humblot, 1989.