

Деопик Д.В.

Исторический контекст-анализ

Применение принципов контент-анализа к текстам ранних исторических повествований (далее – РИП) – погодным записям – позволило обнаружить определенную специфику в наборе лексических элементов и структуре текстов, создающую разновидность контент-анализа – исторический контент-анализ (далее – ИКА). РИП характеризуются: а) своим набором объектов описания, то есть специфическими приоритетными элементами РИП (имен, дат, социальных детерминативов и т.п.; б) очень высокой долей таких элементов в общем объеме текста; в) своими законами распределения в тексте этих элементов. Присущие всем историческим текстам, эти характеристики в наибольшей степени присущи РИП, весьма единообразны для разных эпох и народов, если это – «первые шаги» в записи событий, воспринимавшихся как реально исторические.

Исследование РИП ряда древних и раннесредневековых обществ в разных частях Старого Света позволило уточнить эти характеристики и соответствующую специфику ИКА. Она заключается в: 1) наличии специфических объектов анализов – приоритетных элементов; 2) возникновении своей системы приоритетов в исследовании, формирующих порядок анализа; 3) специфическом группировании частей текста в результате анализа. В итоге получаются характеристики текста, позволяющие отделить РИП от синхронных и одно-культурных им религиозных или литературных текстов или вставок из них в РИП и т.д.

Порядок проведения исторического контекст-анализа РИП, с которого целесообразнее начать его описание, таков: 1) выделяются, в ходе росписи всего текста, одинаковые значимые слова и устанавливается их частота; 2) определяется плотность распределения наиболее массовых их приоритетных элементов на всем тексте; 3) элементы группируются как по смысловому сходству, так и по встречаемости; их группы сводятся в иерархические смысловые структуры; 4) определяется доля групп и структур в объеме текста и их размещение в нем. Затем в тексте выделяются отрезки (обычно это «сгустки» лексики определенных групп или структур), где возрастает число

слов на единицу времени, что разделяет текст большинства РИП на два вида: «общий план» и «крупный план» («сгустки»).

«Общий план» (далее –«ОП») – краткие стандартные по структуре записи об основных для «первых историков» событиях: войнах, посольствах, жизни монархов, внутренней политики и проч.

«Крупным планом» (далее – «КП») может стать событие из большинства таких групп, если резко возрастает: объем информации на единицу времени, частота употребления нетипичной для текста в целом (то есть не распределенной в нем равномерно) и внутренне родственной лексики; тематическое единство на единицу текста (не времени). Отрывок при этом должен быть не менее нескольких фраз. «КП» в разных традициях есть и стандартный и специфический, всегда он легко отличим от «ОП» ; грани между ними резкие, часто маркированные в тексте. «КП» обычно связан с исключительными событиями из жизни монархов, крупными религиозными или военными событиями, причем произошедшими за более короткий отрезок времени и на более ограниченной территории, чем в среднем для такого же объема текста памятника. «ОП» есть всегда, «КП» – обычно; памятники без него –обычно более ранние в данной традиции. Чем позднее историческое повествование –тем более сливаются «ОП» и «КП», но не полностью, так как «КП» – способ выделить явление резким ростом информации о нем.

Исторический контекст-анализ более эффективен в рамках РИП, применительно к более единообразным текстам «ОП», в которых наивысшая в тексте частота и равномерность распределения по тексту: 1) указаний на место и время действия; 2) имен лиц определенных социальных слоев с частотой упоминания, соответствующей иерархии таких слоев; 3) социальных детерминативов и проч. Это – приоритетные элементы РИП.

Массовость упоминаний каждого из таких текстов приоритетных элементов, большинства вводимых исследователем обобщающих понятий для групп родственных слов (значений признака), позволяет надежно выделить элементы и их признаки. (Пример, элемент – «социальная группа», признак – «крупный феодал», значение признака – «князь»). Это относится и к связям между ними, отраженном в достаточно стандартном списке глаголов, также входящих в список приоритетных элементов;

систематизированные связи – уже более сложная характеристика общества, часто не содержащаяся в тексте в явном виде (пример, преобладание определенных отношений между лицами конкретной эпохи). Следующий уровень сложности отражает повышенные частоты простейших сочетаний глаголов и достаточно часто связанных с ними значений признаков элементов («государство X напало на государство Y»). Обилие в РИП таких стереотипных высказываний и близких к ним облегчает сплошную роспись текста, то есть сплошной контент-анализ, разновидностью которого является ИКА. Выделяемые в тексте признаки сопровождаются списком всех их значений – соответствующих слов источника; тем самым обобщение – проверяемо, а машинный анализ идет на уровне лексики естественного языка.

Сплошная роспись текста дает указанные выше характеристики и позволяет выделить скрытую информацию, а также сравнить и разновременные РИП в рамках единой традиции, и РИП разных народов по предпочтению разных групп социальных детерминативов, разных типов пространственных и временных характеристик однородных явлений и т.п.

Характеристика «ОП» конкретного памятника становится базовой для последующих этапов исторического контекст-анализа:

1) отличие от «ОП» – как «КП», так и тестов других видов, включенных в памятник; «кп» характеризуется тем же методом и делятся на группы.

2) Анализ внутренней структуры самого «ОП». Внутри «ОП» проверяется, по частотам слов и их набору, единообразию его частей, выделенным автором текста, устанавливаются более сложные связи между элементами, признаками, их значениями и разными их сочетаниями (через преимущественную взаимовстречаемость). Выявляются группы, связанные с определенным временем, территорией или видом событий; выявляются вставки из «ОП» других памятников. Затем следуют формализованное сопоставление «ОП» в рамках одной культуры и между культурами.

Сходство «ОП» в большинстве РИП обусловлено преимущественным интересом к одним и тем же сторонам реальной, с точки зрения авторов, жизни; представления их об историческом времени всегда стихийно-линейны. В РИП социальная действительность прошлого, как ее видели авторы, подвергалась достаточно единообразному отсеву и сортировке, в текст попадали одинаковые, в целом, категории

фактов, обладающие в тексте высокой частотой и расположенные достаточно равномерно по единой линейной оси времени. Это позволяет выделить исторический контент-анализ прежде всего как преимущественный анализ специфического вида текстов, а в них – специфических элементов и их связей. Полученные результаты позволяют считать его интересной и перспективной областью исследования как нарративных источников, так и массивов документов прошлого.

Опубликовано в Тезисы докладов и сообщений научного совещания «Комплексные методы в исторических исследованиях». М. 1987, с.165-168