

ФОРМАЛИЗАЦИЯ ПРОЦЕССА ПОРОЖДЕНИЯ ТЕКСТА

Первые опыты формального порождения текстов проделаны человеком еще до появления компьютеров. Так, в 1929 году на русском языке был издан роман английского писателя Корнелиуса Крока "Зеленые яблоки". Он выдержал свыше 600 изданий и пользовался большой популярностью. Вот как автор романа описывает процесс его создания (цитируется по [6, с. 324]): "Вы спрашиваете меня, как я создал такой замечательный роман ... Сидя долгие дни в одиночной камере, я был лишен всего: и газет, и книг, и бумаги ... Наконец, надо мною сжалились и дали груду романов без начала и конца и, по-видимому, разных авторов ... И вот тогда я ухватился за произведения неизвестных авторов, я вырывал страницы из разных книг и соединял их в порядке развертывания моего сюжета. Я рад, что явился новатором, открывшим новый способ создания блестящих бестселлеров. Только ножницы и клей!" В качестве "материала" для своих романов Корнелиус Крок использовал произведения Джека Лондона, марка Твена, Стефана Цвейга, Герберта Уэллса и других писателей.

Аналогичный метод применил в своей работе французский журналист Патрик Тевенон, написавший известный роман "А.А." о бедной актрисе. Он специализировался как журналист на рецензиях о спектаклях и потому в качестве "материала" для романа выбрал журнальные и газетные вырезки о театрах и спектаклях [6, с. 324]. Следуя этому же принципу, нечто подобное выполнила позже электронная машина. Она написала за 61 час 30 минут роман под названием "Слепой фараон", опубликованный в США в 1983 году [6, с. 330]. В публикации [1] приводится достаточно много электронных адресов, по которым можно найти компьютерные гипертексты типа "Лоскутное одеяло", "Жидкое стекло", "Anna Karenina goes to Paradise", "РОМАН", "Гостевая книга Буратино", "Робот Сергей Дацюк" и др.

Если говорить о серьезных научных исследованиях, целью которых является изучение принципов организации текстов и роли их отдельных составляющих посредством компьютерного моделирования процесса создания текста, то они подробно описаны в работах [2; 3].

Вместе с тем, до сих пор нет явных, четко сформулированных высказываний о лингвистических процессах порождения текстов нет. Так, С.Д. Кацнельсон по этому поводу пишет так: "Порождающий процесс начинается с определения темы и стратегии сообщения ... Выработка плана и стратегии сообщения не всегда принимает осознанный характер. Наблюдая речь многих людей, можно заметить в ней ненужные повторы, непоследовательность и сбивчивость мысли, незаконченные фрагменты, перескакивание с родной темы на другую и т.д. При всем этом имплицитно в любой речи имеется план и стратегия изложения; ... В процессе порождения речи можно выделить три основных ступни: речеосмыслительную (или семантическую), лексико-морфологическую и фонологическую.

Первая ступень охватывает все собственно семантические процессы, начиная с квантования элементов сознания (знаний) на отдельные пропозиции и кончая формированием глубинных семантико-синтаксических структур. На второй ступени совершаются процессы отбора лексических единиц и грамматических форм, опосредующие переход от семантико-синтаксических структур предшествующей ступени к конкретным предложениям ...

Каждая ступень характеризуется своими особыми структурными единицами и механизмами порождения. В некотором смысле можно утверждать, что каждая ступень имеет свой "словарь" (т.е. набор дискретных элементов), свой "синтаксис" (механизм преобразования структур одного порядка в структуры другого порядка). На первой ступени в функции дискретных единиц выступают представления и понятия, а в функции порождающих механизмов – структуры содержательной валентности. На второй ступени им на смену приходят лексемы и механизмы формальной валентности [4, с. 121].

Если детально рассмотреть точки зрения лингвистов на проблемы порождения текстов, то можно отметить следующее. Большинство из них также проходят к мысли, что в процессе развертывания общей темы текст членится на иерархическую сеть подтем и микротем. Отмечается, что дискретный процесс текстообразования строится на принципе обратной связи, что он сводится к поиску по программе, сложившейся в мозгу автора на основе его субъективно-социального опыта. При этом выбор последующих элементов (слова, словосочетания, высказывания) детерминируется не жесткой цепью следования языковых единиц, а осмысленным выбором ряда элементов из совокупности форм.

Определяя общий алгоритм порождения текстов, исследователи отмечают, что он может рассматриваться как некоторая диспозиция, т.е. процедура, где на каждом шагу предусмотрены различные возможности продолжения и указаны критерии, по которым определяется результат. Все большее число ученых-лингвистов приходят к выводу, что такие процедуры не могут представлять собой жесткую последовательность операций. Они непременно должны включать вероятностные операции, в ходе которых совершается отбор одного из возможных путей продолжения процесса порождения [4, с. 121]. Вместе с тем среди лингвистов нет единого мнения о значимости вероятностных факторов для лингвистических единиц разных уровней. Часть из них считает эти факторы значимыми лишь для лексического уровня [5, с. 57], другие же предполагают, что вероятностные факторы столь же существенны и на более длинных текстовых цепочках.

Литература

- Андреев, А. Литература без бумаги / А. Андреев // Мир Internet. – 1999. – № 1. – С. 76–81.
Болдасов, М.В. Генерация текстов на естественном языке – состояние вопроса и прикладные системы / М.В. Болдасов, Е.Г. Соколова // Научно-техническая информация. Серия 2. Информационные процессы и системы. – 2005. – № 10. – С. 12–23.

- Болдасов, М.В. Генерация текстов на естественном языке – теории, методы, технологии / М.В. Болдасов, Е.Г. Соколова // научно-техническая информация. Серия 2. Информационные процессы и системы. – 2006. – № 7. – С. 1–15.
- Кацнельсон, С.Д. Типология языка и речевое мышление / С.Д. Кацнельсон. – Л.: Наука, 1972.
- Леонтьев, А.А. Психолингвистические единицы и порождение речевого высказывания / А.А. Леонтьев. – М.: Наука, 1969.
- Пекелис, В. Кибернетическая смесь / В. Пекелис. – М.: Знание, 1992.