

СЕМАНТИЧЕСКАЯ СТРУКТУРА ТЕКСТА И СПОСОБЫ ЕЕ ФОРМАЛЬНОГО ПРЕДСТАВЛЕНИЯ

Н. А. Метлицкая

*Белорусский государственный университет, Минск, Республика Беларусь,
bsu@bsu.by*

В статье рассматриваются основные взгляды на семантическую структуру текста и способы ее формального представления.

Ключевые слова: семантическая структура текста; содержание; смысл текста; семантические сети; фреймы.

Текст представляет собой сложное многоуровневое образование, которое одновременно включает в себя языковые, логические, семантические, стилистические, экспрессивные и другие компоненты. При создании системы автоматического порождения текста важным этапом является формальное представление его семантической структуры. Рассмотрим понятие *семантическая структура*.

В современной лингвистике по отношению к структуре текста дифференцируются понятия «семантика», «содержание» и «смысл». Согласно Е.А. Баженовой, «под **семантикой** понимаются значения составляющих текст языковых единиц, обладающих способностью во взаимодействии друг с другом отражать объективную действительность. **Содержание** текста составляют знания о действительности, объективируемые в произведении автором. В свою очередь, **смысл** соотносится со знаниями, получаемыми реципиентом при восприятии текста. Отсюда, **семантическая структура** текста связана с его знаковым характером, **содержательная структура** – с кодированием заключенной в нем информации, **смысловая структура** – с декодированием информации» [2, с. 388].

Рассматривая термины *содержание* и *смысл* более подробно, можно сказать, что *содержание*, как термин семантики текста, относят к информации, заключенной в тексте в целом. В структуре содержания отражаются и замысел автора, и внешняя действительность, и значения слов, словосочетаний, предложений, и фоновые знания автора [4]. Известно, что каждый текст описывает некоторый фрагмент (ситуацию) действительности. При создании текста автор выделяет элементы описываемой ситуации, приписывая им определенные для данной ситуации значения. В итоге содержание текста определяется его денотативной структурой, формируемой в сознании человека; оно реализуется в

тексте подбором конкретных слов и их распределением по тексту. Содержание текста, таким образом, определяется его автором и складывается из значений, непосредственно представленных языковыми знаками.

Смысл текста, по мнению ряда исследователей, представляет собой воспринятую получателем структуру, на основе которой строится модель, представляющая собой авторский замысел. Иными словами, смысл текста можно рассматривать как проекцию содержания на сознание интерпретатора, или интерпретацию того, что сообщается в тексте.

Из вышесказанного можно заключить, что каждый текст имеет только одно содержание, определяемое автором, и может иметь для разных реципиентов различные смыслы.

Существуют различные способы формального представления содержания текста. Одним из таких способов является *семантическая сеть*. Семантическая сеть – это модель предметной области, представленная в виде графа, вершины которого (узлы) соответствуют объектам предметной области, а дуги (рёбра) задают отношения между ними [1]. Объектами могут быть понятия, события, свойства, процессы.

В виде **узлов** часто применяются так называемые семантико-синтаксические функции (семантические падежи) типа АГЕНС, ПАЦИЕНС, ОБЪЕКТ, РЕЦИПИЕНТ и т.д.

В качестве **отношений** между объектами наиболее часто используются следующие: таксономические («класс – подкласс – экземпляр», «множество – элемент»); структурные («часть – целое»); функциональные (определяемые обычно глаголами «производит», «влияет», «умеет» и т.п.); количественные («больше», «меньше», «равно» и т.п.); пространственные («далеко от», «близко от», «за», «под», «над» и т.п.); временные («раньше», «позже», «в течение» и т.п.); атрибутивные («иметь свойство», «иметь значение»); логические (И, ИЛИ, НЕ); казуальные (причинно-следственные).

К достоинствам семантических сетей можно отнести универсальность, достигаемую за счет выбора соответствующего набора отношений; наглядность системы знаний, представленной графически; близость структуры сети, представляющей систему знаний, семантической структуре фраз на естественном языке. Основным недостатком считают тот факт, что при описании систем реального уровня сложности представление, использование и модификация знаний в виде семантических сетей оказывается трудоемкой процедурой, особенно при наличии множественных отношений между ее понятиями.

Еще одним способом формального представления семантической структуры текста является *фреймовый метод*. Фрейм – это модель абстрактного образа, минимально возможное описание сущности какого-либо объекта, явления, события, ситуации, процесса [1].

Фрейм состоит из **имени** и отдельных единиц, называемых **слотами**. Все слоты также получают свои уникальные имена (например, «Количество студентов в группе», «Возраст» и т.д.). Далее следует **тип данных**, который определяет тип значения слота. Среди наиболее употребляемых типов выделяют, например: *frame* – указатель на фрейм; *real* – вещественное число; *integer* – целое число; *boolean* – логический тип; *table* – таблица, *text* – фрагмент текста и т.д. Значение слота должно соответствовать указанному типу данных. Присоединенная процедура, автоматически запускаемая при выполнении некоторого условия, называется **демоном**. Фреймы делят на *фреймы-описания* и *ролевые фреймы*. Первые представляют формализованное описание объектов, явлений; вторые – описание событий, процессов.

Следует отметить, что семантические сети и фреймы более известны в прикладной лингвистике как *формальные способы представления знаний* в памяти компьютера. Однако, представляя собой универсальные и довольно четко структурированные модели описания некоторых фрагментов (ситуаций) действительности, описываемых в тексте, данные модели выступают как способы формализации его смыслового содержания. Рассмотренные способы формализации представляют собой *иерархические модели*, т.е. включают в себя вершины верхнего и нижних уровней.

Существует подход к тексту и его семантической структуре с позиции *ключевых слов*. Согласно данному подходу основное смысловое содержание текста в свернутом виде может быть представлено набором «ключевых» или «опорных» слов. Набор ключевых слов образует своеобразную тематическую сетку текста, которая отражает логику смыслового развертывания данного речевого произведения. В работе [3] среди опорных слов выделяют *главные* и *второстепенные опорные слова* (ГОС и ВОС). Там же предлагается способ определения таких слов в тексте: отнесение опорных слов к главным и второстепенным происходит с учетом *коэффициента семантической важности*, который рассчитывается по определенной формуле.

Таковы основные взгляды на семантическую структуру текста и способы ее формального представления на современном этапе. Проведенный анализ исследований, посвященных семантической структуре текста, позволил обнаружить тесную связь трактовки данного явления с понятиями *содержание* и *смысл* текста. Рассмотренные способы формализации семантической структуры текста имеют свои достоинства и не-

достатки. Выбор способа формализации диктуется потребностями конкретной автоматической системы.

Библиографический список

1. *Анисимов, В.В.* Представление знаний в информационных системах [Электронный ресурс]. – Режим доступа: <https://studfile.net/preview/3075803/page:8/#22> – Дата доступа: 28.09.2024.
2. *Баженова, Е.А.* Смысловая структура текста / Е.А. Баженова // Стилистический энцикл. словарь рус. языка / под ред. М.Н. Кожинной. – 2-е изд. – М.: Флинта; Наука, 2006. – С. 388-392.
3. *Зубов, А.В.* Статический аспект содержания текста и его формальное представление / А.В. Зубов // Уч. записки Тартуского гос. университета. Вып. 745. Квантитативная лингвистика и автоматический анализ текстов. – Тарту, 1986. – С. 75-89.
4. *Плотников, Б.А.* Семиотика текста: параграфемика: учеб. пособие / Б.А. Плотников. – Минск: Вышэйшая школа, 1992. – 190 с.