

СИНТАКСИЧЕСКАЯ ОМОНИМИЯ И НЕКОТОРЫЕ ПОДХОДЫ К ЕЕ РАЗРЕШЕНИЮ ПРИ АВТОМАТИЧЕСКОМ АНАЛИЗЕ ТЕКСТА

Данная проблема встала весьма остро с 1960-х годов с появлением систем автоматической обработки текстов. Пока синтаксический анализ производился людьми, такое явление, как синтаксическая омонимия, не замечали, так как не было должного абстрагирования от смысла, омонимия разрешалась прежде, чем ее можно было заметить.

В случае синтаксической омонимии разница в смыслах словосочетаний или предложений не связана с лексической омонимией. **Предложение или словосочетание является синтаксически омонимичным, если ему можно приписать не менее двух синтаксических структур.** (Следует заметить, что на уровне словосочетания и предложения понятия неоднозначности, многозначности (полисемии) и омонимии имеют один и тот же смысл.) Хотя в некоторых работах можно встретить попытки разграничить вышеуказанные явления [1]. В понятие синтаксической структуры входит выделение минимальных синтаксических единиц (членов предложения) и установление синтаксических связей между ними. С синтаксической омонимией мы сталкиваемся тогда, когда в одном словосочетании или предложении можем по-разному выделить или грамматически проинтерпретировать члены предложения; по-разному установить синтаксические связи между ними.

I. Возможность различных связей возникает, когда член предложения может быть включен в структуру более чем одним способом (имеет более одного потенциального «хозяина»): *Я вернулся из командировки в Минск. Звено Ф. П. Кооля из четырёх человек выпустило 9,5 кубометров бруса.* Такую омонимию называют **стрелочной**, так как разные смыслы конструкции можно отобразить с помощью разных стрелок.

II. Синтаксическая омонимия возникает за счет различной разметки формы или синтаксической связи – **разметочная** омонимия.

● **Различная разметка связи.** Пара «хозяин – слуга» выделяется единственным способом, но связь между ними по-разному можно интерпретировать. Очень распространена **омонимия субъектной и объектной связи**: *перевод Набокова* – «выполненный кем-то пере-

вод произведений Набокова» или «выполненный Набоковым перевод чьих-то произведений». Также возникает **омонимия субъекта и второго объекта**: *Он прислан к нам инструктором* – «инструктор прислал его» или «кто-то прислал его в качестве инструктора».

Является проблемой и полисемия первообразных предлогов. Практически все 20 первообразных русских предлогов многозначны (в разной степени). Например, для предлога *за* выделяются 18 значений, каждое из которых порождает специфическую роль именной группы в синтаксической структуре предложения. Иногда омонимия разных падежных значений может создавать комический эффект: *Беспрерывно заполняются мешки женщинами, мужчины их быстро грузят и увозят в поле.*

- **Различная разметка формы.** Например: *Он не выучил правила* (Р. п. ед. ч. или В. п. мн. ч.?)

- **Одновременная омонимия форм и типов связей.** Различная интерпретация форм может быть связана с различной интерпретацией связей, например: *звонок Марии* (связь адресатная или субъектная? *Марии* – форма Д. п. или Р. п.?)

Нередко в русском языке встречается омонимия предложений типа *Мать любит дочь. Металл вытесняет дерево.*

Существует и **конституентная омонимия**: рядом расположенные слова могут по-разному объединяться в группы (чаще всего из-за неоднозначного выделения однородных членов, реже из-за различного осмысления частиц). *Получили новые тетради и учебники* – непонятно, новые только тетради или учебники тоже.

Рассмотренные типы омонимии могут комбинироваться между собой. Например, в словосочетании *звонок любимой из Москвы* есть разметочная омонимия *звонок любимой* (кому? или чей?) и стрелочная омонимия: «звонок из Москвы» или «любимой из Москвы».

Если синтаксическая омонимия сочетается с лексической, то омонимия называется **лексико-синтаксической**. Например, в словосочетании *простой солдат* можно не только по-разному установить синтаксические связи и интерпретировать форму *солдат*, но и слово *простой* может быть понято и как форма существительного, и как форма прилагательного.

Существует и глубинно-синтаксическая омонимия. Имеется одна и та же синтаксическая структура, но она имеет несколько

смыслов: *Перед докладом на конференции ко мне подошел аспирант.*

Отглагольные существительные со значением действия в русском языке не сохраняют значения глагольного вида и времени. Поэтому неоднозначно предложение *Мы узнали об открытии выставки: Мы узнали, что выставка 1) открылась, 2) открывається, 3) откроется, 4) будет открыта.*

Но при автоматическом анализе текста для разрешения омонимии используются разные подходы. Для иллюстрации используем различные конструкции с предлогом *с*.

1. **Грамматический**, основанный на формальных признаках управляемого или управляющего компонента синтагмы с описываемым предлогом. Но такие признаки можно найти у небольшого числа значений: *ростом с Васю, проживёшь с Васиной*. Но предложения *Пришли учитель с директором* и *Приехали грузовики с мотоциклами* не будут разобраны правильно.

2. **Лексико-грамматический способ**. Когда грамматические критерии существуют, но их оказывается недостаточно и привлекаются и лексические. При различении *отправить с курьером* (посредник) и *уехать с поездом* (средство) для словосочетаний со значением «посредник» управляющие глаголы, как и управляющие существительные, можно задать списком: *послать, передача* и т. д. А управляемые существительные должны быть одушевленными (исключение *оказия* и *почта*). Для словосочетаний со значением «средство» управляемые существительные всегда неодушевленные и составляют лексический класс «средства сообщений», а управляющие глаголы – глаголы движения.

3. **Лексический способ**. Используется при невозможности привлечения грамматических критериев. К таким критериям относятся вхождение управляющего (управляемого) члена в определенный лексико-семантический класс, возможность задать списком члены синтагмы и комбинация этих критериев. Например, в словосочетании *занят с утра* («начало действия») существительное относится к лексическому классу «время».

Но всё-таки далеко не все различимые человеком значения предлога *с* могут быть выделены автоматически при помощи описанных выше способов. В автоматизированной системе, где не используется метод контроля, на предсинтаксическом этапе анализа

только машина делает выбор в пользу интерпретации конструкции тем или иным способом (способами).

1. *Соколова Е. Г.* Об организации формализованного синтаксического представления в терминах членов предложения // Машинный перевод и прикладная лингвистика: Сб. науч. тр. М., 1986. Вып. 271.