

СНЯТИЕ ЧАСТЕРЕЧНОЙ НЕОДНОЗНАЧНОСТИ  
В АНГЛИЙСКОМ ЯЗЫКЕ КОРПУСНЫМИ МЕТОДАМИ  
(на примере лексемы *CALL*)

Частеречная неоднозначность, являясь одной из ключевых проблем обработки естественного языка, и ее разрешение коренятся в динамичности и сложности языковой системы. Традиционно различают 3 типа неоднозначности: частеречная неоднозначность, морфологическая неоднозначность, лексическая неоднозначность (К. К. Боярский, Е. А. Каневский, С. К. Стафеев, 2012). Для каждого из этих типов неоднозначности будут использоваться свои методы снятия.

В рамках нашего исследования мы анализировали способы разрешения частеречной неоднозначности лексемы *call* в контексте с помощью формулирования грамматических правил. Для создания правил использовалась программа NooJ. С её помощью пользователи могут создавать грамматики, используя либо текстовый редактор или графический редактор. В ходе эксперимента мы проанализировали лексему *call* в произведении «Звонок» Грейс Макгоуэн Кук. Было обнаружено 20 случаев частеречной неоднозначности. Для снятия частеречной неоднозначности лексемы *call* необходимо проанализировать её окружение и задать правила для проверки. Лексема *call* может выступать в роли существительного или глагола в зависимости от контекста. Так, словоформа *call* является глаголом, если выполняется один из следующих признаков: наличие прямого дополнения (например, *call a friend*), наличие наречий после словоформы, наличие вспомогательных глаголов, наличие частицы *to* перед словоформой, а также наличие отрицаний. С другой стороны, словоформа *call* является существительным, если перед ней стоят артикли *a*, *an*, *the*, прилагательные, притяжательные местоимения или предлоги. Далее с помощью программы NooJ мы создали грамматики для снятия частеречной неоднозначности. В результате применения грамматик удалось решить проблему частеречной неоднозначности в 18 случаях из 20. В оставшихся случаях затруднилось из-за намеренного нарушения автором орфографии и грамматики английского языка.

Таким образом, на примере лексемы *call* был рассмотрен эффективный способ снятия частеречной неоднозначности, основанный на создании грамматических правил. Этот подход позволяет определить частеречную принадлежность словоформы, опираясь исключительно на её контекстное окружение. Такой метод демонстрирует свою универсальность и применимость для

обработки естественного языка. Это особенно актуально для задач автоматической обработки текстов, где важно минимизировать зависимость от семантического анализа и максимизировать опору на формальные признаки.